

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 18:02:37 ; Search time 23 Seconds  
(without alignments)  
2666.594 Million cell updates/sec

Title: US-09-980-403-2  
Perfect score: 1188  
Sequence: 1 MDLPRGLVAVALLSLWPGET.....FRSARRRRFGLDTPKVLK 1188

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	1.8	158	4	US-09-312-283C-391
2	17	1.4	85	4	US-09-312-283C-390
3	15	1.3	214	4	US-10-061-658-5
4	15	1.3	214	4	US-10-061-658-6
5	15	1.3	214	4	US-10-061-658-9
6	14	1.2	1183	4	US-09-532-310B-5
7	11	0.9	122	6	5196511-9
8	11	0.9	1183	4	US-09-532-310B-6
9	9	0.8	9	1	US-08-462-128-34
10	9	0.8	9	1	US-08-463-180-34
11	8	0.7	17	1	US-08-173-497-20
12	8	0.7	17	1	US-08-236-889-20
13	8	0.7	17	1	US-08-485-618-20
14	8	0.7	17	1	US-08-362-652-20
15	8	0.7	17	2	US-08-605-672-20
16	8	0.7	17	2	US-08-482-293A-20
17	8	0.7	17	2	US-08-943-363-20
18	8	0.7	17	3	US-09-193-043-20
19	8	0.7	17	4	US-09-688-307A-20
20	8	0.7	17	4	US-09-350-259-20
21	8	0.7	76	1	US-08-848-652-4
22	8	0.7	76	3	US-09-083-521-5
23	8	0.7	80	1	US-08-848-252-2
24	8	0.7	95	1	US-07-626-618A-2
25	8	0.7	95	1	US-07-928-611-2
26	8	0.7	95	1	US-08-333-977-2
27	8	0.7	95	2	US-08-487-811A-2
28	8	0.7	95	2	US-09-060-694-2
29	8	0.7	95	2	US-09-378-074-2
30	8	0.7	95	2	PCT-US93-07370-2
31	8	0.7	97	4	US-08-973-544-4
32	8	0.7	104	4	US-08-973-544-3
33	8	0.7	112	4	US-09-489-039A-10236
34	8	0.7	179	4	US-09-419-568F-28
35	8	0.7	179	4	US-09-354-243B-28
36	8	0.7	179	4	US-09-870-574-2
37	8	0.7	183	4	US-09-252-991A-25905
38	8	0.7	234	3	US-09-040-483-5
39	8	0.7	234	4	US-08-740-036-5
40	8	0.7	315	1	US-08-118-270-28
41	8	0.7	315	5	PCT-US93-08528-48
42	8	0.7	345	4	US-09-489-039A-7417
43	8	0.7	374	3	US-09-046-736-4
44	8	0.7	387	1	US-07-626-618A-17
45	8	0.7	387	1	US-08-086-438C-3
46	8	0.7	387	1	US-08-056-051-2
47	8	0.7	387	1	US-07-928-611-18
48	8	0.7	387	1	US-08-333-977-17
49	8	0.7	387	1	US-08-449-153-1
50	8	0.7	387	1	US-08-434-877-3
51	8	0.7	387	1	US-08-449-651-1
52	8	0.7	387	2	US-08-487-811A-18
53	8	0.7	387	3	US-08-833-703-1
54	8	0.7	387	3	US-08-475-742-4
55	8	0.7	387	3	US-09-060-694-18
56	8	0.7	387	4	US-09-378-074-18
57	8	0.7	387	4	US-08-261-293-4
58	8	0.7	387	5	PCT-US93-07370-18
59	8	0.7	419	1	US-08-056-051-4
60	8	0.7	419	1	US-07-928-611-20
61	8	0.7	419	2	US-08-487-811A-20
62	8	0.7	419	3	US-09-060-694-20
63	8	0.7	419	4	US-09-378-074-20
64	8	0.7	419	5	PCT-US93-07370-20
65	8	0.7	453	4	US-09-252-991A-17394
66	8	0.7	467	1	US-08-056-051-6
67	8	0.7	467	1	US-07-928-611-22
68	8	0.7	467	2	US-08-487-811A-22
69	8	0.7	467	3	US-09-046-736-2
70	8	0.7	467	3	US-09-060-694-22
71	8	0.7	467	4	US-09-378-074-22
72	8	0.7	467	5	PCT-US93-07370-22
73	8	0.7	494	1	US-08-485-618-103
74	8	0.7	494	2	US-08-605-672-103
75	8	0.7	494	2	US-08-482-293A-103
76	8	0.7	494	2	US-08-943-363-103
77	8	0.7	494	3	US-09-193-043-103
78	8	0.7	494	4	US-09-688-307A-103
79	8	0.7	494	4	US-09-350-259-103
80	8	0.7	605	2	US-08-472-666-1
81	8	0.7	605	5	PCT-US96-07615-1
82	8	0.7	628	4	US-09-252-991A-37277
83	8	0.7	877	4	US-09-252-991A-25547
84	8	0.7	1151	1	US-08-286-889-37
85	8	0.7	1151	1	US-08-485-618-37
86	8	0.7	1151	2	US-08-362-652-37
87	8	0.7	1151	2	US-08-605-672-37
88	8	0.7	1151	2	US-08-482-293A-37
89	8	0.7	1151	2	US-08-943-363-37
90	8	0.7	1151	3	US-09-193-043-37
91	8	0.7	1151	4	US-09-688-307A-37
92	8	0.7	1151	4	US-09-350-259-37
93	8	0.7	1152	2	US-08-476-062A-43
94	8	0.7	1152	5	PCT-US96-01314-43
95	8	0.7	1152	6	542439-2
96	8	0.7	1153	1	US-08-173-497-3
97	8	0.7	1153	1	US-08-286-889-3
98	8	0.7	1153	1	US-08-485-618-3
99	8	0.7	1153	1	US-08-362-652-3
100	8	0.7	1153	2	US-08-605-672-3

Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 4, Appli  
Sequence 3, Appli  
Sequence 10236, A  
Sequence 28, Appli  
Sequence 28, Appli  
Sequence 2, Appli  
Sequence 25905, A  
Sequence 5, Appli  
Sequence 5, Appli  
Sequence 28, Appli  
Sequence 28, Appli  
Sequence 7417, Ap  
Sequence 4, Appli  
Sequence 3, Appli  
Sequence 2, Appli  
Sequence 17, Appli  
Sequence 18, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 18, Appli  
Sequence 18, Appli  
Sequence 4, Appli  
Sequence 18, Appli  
Sequence 4, Appli  
Sequence 20, Appli  
Sequence 20, Appli  
Sequence 20, Appli  
Sequence 20, Appli  
Sequence 17394, A  
Sequence 6, Appli  
Sequence 22, Appli  
Sequence 22, Appli  
Sequence 22, Appli  
Sequence 22, Appli  
Sequence 103, App  
Sequence 103, App  
Sequence 103, App  
Sequence 103, App  
Sequence 103, App  
Sequence 103, App  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 32727, A  
Sequence 25547, A  
Sequence 37, Appli  
Sequence 37, Appli  
Sequence 37, Appli  
Sequence 37, Appli  
Sequence 37, Appli  
Sequence 37, Appli  
Sequence 37, Appli  
Sequence 37, Appli  
Sequence 37, Appli  
Sequence 37, Appli  
Sequence 37, Appli  
Sequence 43, Appli  
Sequence 43, Appli  
Patent No. 5424399  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 3, Appli

## ALIGNMENTS

## RESULT 1

US-09-312-283C-391

; Sequence 391, Application US/09312283C

; Patent No. 6573095

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Murison, James G.

; APPLICANT: Kumble, Krishanand D.

; TITLE OF INVENTION: Compositions isolated from Skin Cells

; FILE OF INVENTION: and Methods for Their Use

; FILE REFERENCE: 11000.1011c2

; CURRENT APPLICATION NUMBER: US/09/312,283C

; CURRENT FILING DATE: 1999-05-14

; NUMBER OF SEQ ID NOS: 425

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 391

; LENGTH: 158

; TYPE: PRT

; ORGANISM: Mouse

US-09-312-283C-391

Query Match 1.8%; Score 21; DB 4; Length 158;

Best Local Similarity 100.0%; Pred. No. 1e-11;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

636 RPVVQINASHFEPSKINIPH 656

|||||

57 RPVVQINASHFEPSKINIPH 77

DB

US-09-312-283C-390

; Sequence 390, Application US/09312283C

; Patent No. 6573095

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Murison, James G.

; APPLICANT: Kumble, Krishanand D.

; TITLE OF INVENTION: Compositions Isolated from Skin Cells

; FILE OF INVENTION: and Methods for Their Use

; FILE REFERENCE: 11000.1011c2

; CURRENT APPLICATION NUMBER: US/09/312,283C

; CURRENT FILING DATE: 1999-05-14

; NUMBER OF SEQ ID NOS: 425

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 390

; LENGTH: 85

; TYPE: PRT

; ORGANISM: Mouse

US-09-312-283C-390

Query Match 1.4%; Score 17; DB 4; Length 85;

Best Local Similarity 100.0%; Pred. No. 3.7e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

528 NGTLKDSHSYQNARTGS 544

|||||

28 NGTLKDSHSYQNARTGS 44

DB

US-10-061-658-5

RESULT 3

US-10-061-658-5

; Sequence 5, Application US/10061658

; Patent No. 6652856

; GENERAL INFORMATION:

; APPLICANT: Biogen, Inc.

; APPLICANT: Gotwals, Philip

; APPLICANT: Kotliansky, Victor

; TITLE OF INVENTION: Method for the Treatment of Fibrosis

; FILE REFERENCE: A073US

; CURRENT APPLICATION NUMBER: US/10/061,658

; CURRENT FILING DATE: 2002-02-01

; PRIOR APPLICATION NUMBER: 60/137,214

; PRIOR FILING DATE: 1999-06-01

; PRIOR APPLICATION NUMBER: 60/130,847

; PRIOR FILING DATE: 1999-04-22

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 214

; TYPE: PRT

; ORGANISM: Rat

US-10-061-658-5

Query Match 1.3%; Score 15; DB 4; Length 214;

Best Local Similarity 100.0%; Pred. No. 6.6e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

164 DIVIVLDGNSIYPW 178

|||||

22 DIVIVLDGNSIYPW 36

DB

RESULT 4

US-10-061-658-6

; Sequence 6, Application US/10061658

; Patent No. 6652856

; GENERAL INFORMATION:

; APPLICANT: Biogen, Inc.

; APPLICANT: Gotwals, Philip

; APPLICANT: Kotliansky, Victor

; TITLE OF INVENTION: Method for the Treatment of Fibrosis

; FILE REFERENCE: A073US

; CURRENT APPLICATION NUMBER: US/10/061,658

; CURRENT FILING DATE: 2002-02-01

; PRIOR APPLICATION NUMBER: 60/137,214

; PRIOR FILING DATE: 1999-06-01

; PRIOR APPLICATION NUMBER: 60/130,847

; PRIOR FILING DATE: 1999-04-22

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 214

; TYPE: PRT

; ORGANISM: Homo sapien

US-10-061-658-6

Query Match 1.3%; Score 15; DB 4; Length 214;

Best Local Similarity 100.0%; Pred. No. 6.6e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

164 DIVIVLDGNSIYPW 178

|||||

22 DIVIVLDGNSIYPW 36

DB

RESULT 5

US-10-061-658-9

; Sequence 9, Application US/10061658

; Patent No. 6652856

; GENERAL INFORMATION:

; APPLICANT: Biogen, Inc.

; APPLICANT: Gotwals, Philip

; APPLICANT: Kotliansky, Victor

; TITLE OF INVENTION: Method for the Treatment of Fibrosis

FILE REFERENCE: A073US  
CURRENT APPLICATION NUMBER: US/10/061,658  
CURRENT FILING DATE: 2002-02-01  
PRIOR APPLICATION NUMBER: 60/137,214  
PRIOR FILING DATE: 1999-06-01  
PRIOR APPLICATION NUMBER: 60/130,847  
PRIOR FILING DATE: 1999-04-22  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 214  
TYPE: PRT  
ORGANISM: Homo sapien  
US-10-061-658-9

Query Match 1.3%; Score 15; DB 4; Length 214;  
Best Local Similarity 100.0%; Pred. No. 6.6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178  
DB 22 DIVVLGSGNSIYPW 36

## RESULT 6

US-09-532-310B-5  
Sequence 5, Application US/09532310B

Patent No. 6596276

GENERAL INFORMATION:

APPLICANT: Senger, Donald R

Detmar, Michael

Claffey, Kevin P

TITLE OF INVENTION: Method for inhibiting tumor

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: David Prashker, Esq.

STREET: P.O. Box 5387

CITY: Magnolia

STATE: Massachusetts

COUNTRY: USA

ZIP: 01930

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.40 Mb storage

COMPUTER: Dell PC

OPERATING SYSTEM: MS DOS

SOFTWARE: Microsoft Word version 97

CURRENT APPLICATION DATA: US/09/532,310B

FILING DATE: 22-Mar-2000

CLASSIFICATION: Unknown

ATTORNEY/AGENT INFORMATION:

NAME: David Prashker, Esq.

REGISTRATION/DOCKET NUMBER: BIS-036

TELECOMMUNICATION INFORMATION:

TELEPHONE: (978) 525-3794

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1183 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-532-310B-5

Query Match 1.2%; Score 14; DB 4; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 0.00027;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYP 177  
|||||

Db 174 DIVVLGSGNSIYP 187

## RESULT 7

5196511-9

Patent No. 5196511

APPLICANT: PLOW, EDWARD F.; D'SOUZA, STANLEY E.

TITLE OF INVENTION: PEPTIDES AND ANTIBODIES THAT INHIBIT

INTEGRIN-LIGAND BINDING

NUMBER OF SEQUENCES: 31

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/444,777

FILING DATE: 01-DEC-1989

SEQ ID NO: 9:

LENGTH: 22

5196511-9

Query Match

Best Local Similarity 0.9%; Score 11; DB 6; Length 22;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAPMY 505

DB 11 TDVLLVGAPMY 21

## RESULT 8

US-09-532-310B-6

Sequence 6, Application US/09532310B

Patent No. 6596276

GENERAL INFORMATION:

APPLICANT: Senger, Donald R

Detmar, Michael

Claffey, Kevin P

TITLE OF INVENTION: Method for inhibiting tumor

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: David Prashker, Esq.

STREET: P.O. Box 5387

CITY: Magnolia

STATE: Massachusetts

COUNTRY: USA

ZIP: 01930

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.40 Mb storage

COMPUTER: Dell PC

OPERATING SYSTEM: MS DOS

SOFTWARE: Microsoft Word version 97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/532,310B

FILING DATE: 22-Mar-2000

CLASSIFICATION: Unknown

ATTORNEY/AGENT INFORMATION:

NAME: David Prashker, Esq.

REGISTRATION NUMBER: 29,693

REFERENCE/DOCKET NUMBER: BIS-036

TELECOMMUNICATION INFORMATION:

TELEPHONE: (978) 525-3794

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 1183 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-532-310B-6

Query Match 0.9%; Score 11; DB 4; Length 1183;  
Best Local Similarity 100.0%; Pred. No. 0.19;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 495 TDVLLVGAPMY 505  
 Db 506 TDVLLVGAPMY 516

## RESULT 9

US-08-462-128-34  
 ; Sequence 34, Application US/08462128  
 ; Patent No. 5686059  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Goetinck, Paul F.  
 ; APPLICANT: Tondravi, Mehرداد  
 ; TITLE OF INVENTION: CARTILAGE MATRIX PROTEIN AND METHODS FOR  
 ; TITLE OF INVENTION: USE  
 ; NUMBER OF SEQUENCES: 43  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Lahive & Cockfield  
 ; STREET: 60 State Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/462,128  
 ; FILING DATE:  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/006,096  
 ; FILING DATE: 15-JAN-1993  
 ; APPLICATION NUMBER: US 07/866,403  
 ; FILING DATE: 10-APR-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Myers, Paul L.  
 ; REGISTRATION NUMBER: 35,965  
 ; REFERENCE/DOCKET NUMBER: MGP-005  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5941  
 ; INFORMATION FOR SEQ ID NO: 34:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 9 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear

US-08-462-128-34

Query Match 0.8%; Score 9; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 164 DIVIVLDGS 172  
 Db 1 DIVIVLDGS 9

## RESULT 10

US-08-463-180-34  
 ; Sequence 34, Application US/08463180  
 ; Patent No. 5741670  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Goetinck, Paul F.  
 ; APPLICANT: Tondravi, Mehرداد  
 ; TITLE OF INVENTION: CARTILAGE MATRIX PROTEIN AND METHODS FOR  
 ; TITLE OF INVENTION: USE  
 ; NUMBER OF SEQUENCES: 43  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Lahive & Cockfield  
 ; STREET: 60 State Street

CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/463,180  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/006,096  
 ; FILING DATE: 15-JAN-1993  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/866,403  
 ; FILING DATE: 10-APR-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Myers, Paul L.  
 ; REGISTRATION NUMBER: 35,965  
 ; REFERENCE/DOCKET NUMBER: MGP-005DV  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5941  
 ; INFORMATION FOR SEQ ID NO: 34:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 9 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; US-08-463-180-34

Query Match 0.8%; Score 9; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVIVLDGS 172  
 Db 1 DIVIVLDGS 9

## RESULT 11

US-08-173-497-20  
 ; Sequence 20, Application US/08173497  
 ; Patent No. 5437958  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gallatin, W. Michael  
 ; APPLICANT: Van Der Vieren, Monica  
 ; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha  
 ; TITLE OF INVENTION: Subunit  
 ; NUMBER OF SEQUENCES: 29  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 ; STREET: 233 S. Wacker Drive, 6300 Sears Tower  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606-6402  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/173,497  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. 5437958and, Greta E.  
 ; REGISTRATION NUMBER: 35,302  
 ; REFERENCE/DOCKET NUMBER: 27866/31363



TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-173-497-20

Query Match 0.7%; Score 8; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 12  
US-08-286-889-20  
; Sequence 20, Application US/08286889  
; Patent No. 5470953  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Mich  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-286-889-20

Query Match 0.7%; Score 8; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 13  
US-08-485-618-20  
; Sequence 20, Application US/08485618  
; Patent No. 5728533  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,618  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-485-618-20

Query Match 0.7%; Score 8; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 14  
US-08-362-652-20  
; Sequence 20, Application US/08362652  
; Patent No. 5766850  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago

STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/362,652  
FILING DATE:  
CLASSIFICATION: 430  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32391  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-362-652-20

Query Match 0.7%; Score 8; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 15  
US-08-605-672-20  
Sequence 20, Application US/0805672  
Patent No. 5817515  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/605,672  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889

FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32684  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-605-672-20

Query Match 0.7%; Score 8; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 16  
US-08-482-293A-20  
Sequence 20, Application US/08482293A  
Patent No. 5831029  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,293A  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32684  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-482-293A-20

Query Match 0.7%; Score 8; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 17

US-08-943-363-20  
Sequence 20, Application US/08943363  
Patent No. 5837478  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/943,363  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32684  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-943-363-20

Query Match 0.7%; Score 8; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 18  
US-09-193-043-20  
Sequence 20, Application US/09193043  
Patent No. 6251395  
GENERAL INFORMATION:  
APPLICANT: Gallatin, Michael W.  
TITLE OF INVENTION: No. 6251395el Human 2  
FILE REFERENCE: 27866/35004  
CURRENT APPLICATION NUMBER: US/09/193,043  
CURRENT FILING DATE: 1998-11-16  
EARLIER APPLICATION NUMBER: 08/173,497  
EARLIER FILING DATE: 1993-12-23  
EARLIER APPLICATION NUMBER: 08/286,889  
EARLIER FILING DATE: 1994-08-05  
EARLIER APPLICATION NUMBER: 08/362,652  
EARLIER FILING DATE: 1994-12-21  
EARLIER APPLICATION NUMBER: 08/943,363  
EARLIER FILING DATE: 1997-10-03  
NUMBER OF SEQ ID NOS: 114  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 20  
LENGTH: 17  
TYPE: PRT  
ORGANISM: dog  
US-09-193-043-20

Query Match 0.7%; Score 8; DB 3; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
DB 1 LVVGAPLE 8

RESULT 19

US-09-688-307A-20  
Sequence 20, Application US/09688307A  
Patent No. 6432404  
GENERAL INFORMATION:  
APPLICANT: Gallatin, Michael W.  
TITLE OF INVENTION: No. 6432404el Human Beta-2  
FILE REFERENCE: 27866/36646  
CURRENT APPLICATION NUMBER: US/09/688,307A  
CURRENT FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 09/193,043  
PRIOR FILING DATE: 1998-11-16  
PRIOR APPLICATION NUMBER: 08/605,672  
PRIOR FILING DATE: 1996-02-22  
PRIOR APPLICATION NUMBER: 08/173,497  
PRIOR FILING DATE: 1993-12-23  
PRIOR APPLICATION NUMBER: 08/286,889  
PRIOR FILING DATE: 1994-08-05  
PRIOR APPLICATION NUMBER: 08/362,652  
PRIOR FILING DATE: 1994-12-21  
PRIOR APPLICATION NUMBER: 08/943,363  
PRIOR FILING DATE: 1997-10-03  
NUMBER OF SEQ ID NOS: 114  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 20  
LENGTH: 17  
TYPE: PRT  
ORGANISM: dog  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 13  
OTHER INFORMATION: Xaa = any or unknown amino acid  
US-09-688-307A-20

Query Match 0.7%; Score 8; DB 4; Length 17;

Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
Db 1 LVVGAPLE 8

## RESULT 20

US-09-350-259-20  
Sequence 20, Application US/09350259  
Patent No. 6620915  
GENERAL INFORMATION:  
APPLICANT: Gallatin, Michael W.  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 6620915el Human 2  
FILE REFERENCE: 27866/35004  
CURRENT APPLICATION NUMBER: US/09/350,259  
CURRENT FILING DATE: 1999-07-08  
EARLIER APPLICATION NUMBER: 09/193,043  
EARLIER FILING DATE: 1998-11-16  
EARLIER APPLICATION NUMBER: 08/173,497  
EARLIER FILING DATE: 1993-12-23  
EARLIER APPLICATION NUMBER: 08/286,889  
EARLIER FILING DATE: 1994-08-05  
EARLIER APPLICATION NUMBER: 08/362,652  
EARLIER FILING DATE: 1994-12-21  
EARLIER APPLICATION NUMBER: 08/943,363  
EARLIER FILING DATE: 1997-10-03  
NUMBER OF SEQ ID NOS: 114  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 20  
LENGTH: 17  
TYPE: PRT  
ORGANISM: dog  
US-09-350-259-20

Query Match 0.7%; Score 8; DB 4; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
Db 1 LVVGAPLE 8

## RESULT 21

US-08-848-252-4  
Sequence 4, Application US/08848252  
Patent No. 5804177  
GENERAL INFORMATION:  
APPLICANT: Humphries, Keith R.  
TITLE OF INVENTION: METHOD OF USING CD24 AS A CELL MARKER  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bereskin & Parr  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/848,252  
FILING DATE: 29-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/538,052  
FILING DATE:

APPLICATION NUMBER: US 08/151,672  
FILING DATE: 15-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: McDiarmid, Shona S.  
REGISTRATION NUMBER: P-38,798  
REFERENCE/DOCKET NUMBER: 3158-028  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
TELEX: 06-23115  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 76 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-848-252-4

Query Match 0.7%; Score 8; DB 1; Length 76;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLALL 1159  
Db 12 GLLALL 19

## RESULT 22

US-09-083-521-5  
Sequence 5, Application US/09083521  
Patent No. 6048970  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/083,521  
FILING DATE: Herewith  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: CERRONE, MICHAEL C.  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0527 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 76 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1216498  
US-09-083-521-5

Query Match 0.7%; Score 8; DB 3; Length 76;  
Best Local Similarity 100.0%; Pred. No. 11;

```
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 12 GLLLLALL 19

RESULT 23
US-08-848-252-2
; Sequence 2, Application US/08848252
; Patent No. 5804177
; GENERAL INFORMATION:
; APPLICANT: Humphries, Keith R.
; TITLE OF INVENTION: METHOD OF USING CD24 AS A CELL MARKER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/848,252
; FILING DATE: 29-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/538,052
; FILING DATE:
; APPLICATION NUMBER: US 08/151,672
; FILING DATE: 15-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McDiarmid, Shona S.
; REGISTRATION NUMBER: P-38,798
; REFERENCE/DOCKET NUMBER: 3158-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; TELEX: 06-23315
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-848-252-2

Query Match 0.7%; Score 8; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 12 GLLLLALL 19

RESULT 24
US-07-626-618A-2
; Sequence 2, Application US/07626618A
; Patent No. 5422265
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civeilli, Olivier
; TITLE OF INVENTION: A No. 5422265el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
```

```
QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 25
US-07-928-611-2
; Sequence 2, Application US/07928611
; Patent No. 5559601
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civeilli, Olivier
; TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/928,611
; FILING DATE: 19920810
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 5569601nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-626-618A-2

Query Match 0.7%; Score 8; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88
```

```

; LENGTH: 95 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-928-611-2
Query Match 0.7%; Score 8; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 26
US-08-333-977-2
; Sequence 2, Application US/08333977
; Patent No. 5594108
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5594108el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/333,977
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/626,618
; FILING DATE: 7 DEC 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5594108nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-333-977-2

Query Match 0.7%; Score 8; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 27
US-08-487-811A-2
; Sequence 2, Application US/08487811A
; Patent No. 5883226
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/060,694
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6203998nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-487-811A-2

Query Match 0.7%; Score 8; DB 2; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 28
US-09-060-694-2
; Sequence 2, Application US/09060694
; Patent No. 6203998
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 6203998el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/060,694
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6203998nan, Kevin E
; REGISTRATION NUMBER: 35,303
```

REFERENCE/DOCKET NUMBER: 90,1092-MM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 95 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-060-694-2

Query Match 0.7%; Score 8; DB 3; Length 95;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 29  
US-09-378-074-2  
Sequence 2, Application US/09378074  
Patent No. 6437114  
GENERAL INFORMATION:  
APPLICANT: Van Tol, Hubert H.M.  
Civelli, Olivier  
TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Allegretti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/378,074  
FILING DATE: 20-Aug-1999  
CLASSIFICATION: <Unknown>  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 07/928,611  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6437114an, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1234  
TELEFAX: 312-715-1234  
TELEX: 810-221-8317  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 95 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-378-074-2

Query Match 0.7%; Score 8; DB 4; Length 95;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

DB 81 LLLALLVL 88

RESULT 30  
PCT-US93-07370-2  
Sequence 2, Application PC/TUS9307370  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/07370  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 95 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US93-07370-2

Query Match 0.7%; Score 8; DB 5; Length 95;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 31  
US-08-973-544-4  
Sequence 4, Application US/08973544  
Patent No. 6338950  
GENERAL INFORMATION:  
APPLICANT: WEISS, Elisabeth  
TITLE OF INVENTION: NEW IMMUNOREGULATORY PROTEIN LST-1  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM LLP  
STREET: 655 Fifteenth St., NW, Suite 300, G St. Lobby  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/973,544  
FILING DATE: 18-DEC-1997  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: PCT EP 96/02663  
FILING DATE: 20-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 95109511.6  
FILING DATE: 20-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 95112201.9  
FILING DATE: 03-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Kitts, Monica Chin  
REGISTRATION NUMBER: 36,105  
REFERENCE/DOCKET NUMBER: P8341-7073  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-5000

```

; TELEFAX: (202) 638-4810
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-973-544-4
Query Match 0.7%; Score 8; DB 4; Length 97;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGILLLA 1157
Db 16 LGGILLLA 23

RESULT 32
US-08-973-544-3
; Sequence 3, Application US/08973544
; Patent No. 6338950
; GENERAL INFORMATION:
; APPLICANT: WEISS, Elisabeth
; TITLE OF INVENTION: NEW IMMUNOREGULATORY PROTEIN LST-1
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAI DO, MARMEI STEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth St., NW, Suite 300, G St. Lobby
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,544
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT EP 96/02663
; FILING DATE: 20-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95109511.6
; FILING DATE: 20-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95112201.9
; FILING DATE: 03-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitts, Morica Chin
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: P8341-7073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-5000
; TELEFAX: (202) 638-4810
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-973-544-3
Query Match 0.7%; Score 8; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGILLLA 1157
Db 16 LGGILLLA 23

us-09-980-403-2.oligo.ra1
; TELEFAX: (202) 638-4810
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-489-039A-10236
Query Match 0.7%; Score 8; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
Db 13 LLLALLLV 20

RESULT 33
US-09-489-039A-10236
; Sequence 10236, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-2004C01
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10236
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10236
Query Match 0.7%; Score 8; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
Db 13 LLLALLLV 20

RESULT 34
US-09-419-568F-28
; Sequence 28, Application US/09419568F
; Patent No. 6331613
; GENERAL INFORMATION:
; APPLICANT: Dumoutier, Laure
; APPLICANT: Louhed, Jamila
; APPLICANT: Renauld, Jean-Christophe
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible
; FILE REFERENCE: LUD 5543.2
; CURRENT APPLICATION NUMBER: US/09/419,568F
; CURRENT FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US09/354,243
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: US09/178,973
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 29
; SEQ ID NO 28
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-419-568F-28
Query Match 0.7%; Score 8; DB 4; Length 179;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
Db 21 LLLALLLV 28

RESULT 35
US-09-354-243B-28
; Sequence 28, Application US/09354243B
; Patent No. 6359117
; GENERAL INFORMATION:
; APPLICANT: Dumoutier, Laure
; APPLICANT: Louhed, Jamila
US-09-354-243B-28
Query Match 0.7%; Score 8; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGILLLA 1157
Db 16 LGGILLLA 23
```



```
; APPLICANT: Renaud, Jean-Christophe
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fa
; TITLE OF INVENTION: (TIFFS)
; TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof
; FILE REFERENCE: LUD 5543.1
; CURRENT APPLICATION NUMBER: US/09/354,243B
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: US09/178,973
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 29
; SEQ ID NO 28
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-354-243B-28

Query Match          0.7%; Score 8; DB 4; Length 179;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160
DB 21 LLLLLLV 28

RESULT 36
US-09-870-574-2
; Sequence 2, Application US/09870574
; Patent No. 6551799
; GENERAL INFORMATION:
; APPLICANT: Gurney, Austin L.
; APPLICANT: Aggarwal, Sudeepa
; APPLICANT: Xie, Ming-Hong
; APPLICANT: Maruoka, Ellen M.
; APPLICANT: Foster, Jessica S.
; APPLICANT: Goddard, Audrey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: INTERLEUKIN-22 POLYPEPTIDES, NUCLEIC ACIDS ENCODING
; FILE REFERENCE: P2806-1(US)
; CURRENT APPLICATION NUMBER: US/09/870,574
; CURRENT FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: US 60/169,495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 7
; SEQ ID NO 2
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-870-574-2

Query Match          0.7%; Score 8; DB 4; Length 179;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160
DB 21 LLLLLLV 28

RESULT 37
US-09-252-991A-25905
; Sequence 25905, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
```

```
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25905
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25905

Query Match          0.7%; Score 8; DB 4; Length 183;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1171 SARRRREP 1178
DB 21 SARRRREP 28

RESULT 38
US-09-040-483-5
; Sequence 5, Application US/09040483
; Patent No. 6143867
; GENERAL INFORMATION:
; APPLICANT: Akerd, Ingrid E.
; TITLE OF INVENTION: NOVEL HUMAN EOSINOPHIL-
; TITLE OF INVENTION: DERIVED BASIC PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,483
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/740,036
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0133 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 544241
US-09-040-483-5

Query Match          0.7%; Score 8; DB 3; Length 234;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1153 LLLALLV 1160  
Db 4 LLLALLV 11

RESULT 39  
US-08-740-036-5  
; Sequence 5, Application US/08740036  
; Patent No. 6492507  
; GENERAL INFORMATION:  
; APPLICANT: Akribiom, Ingrid E.  
; TITLE OF INVENTION: NOVEL HUMAN EOSINOPHIL-  
; TITLE OF INVENTION: DERIVED BASIC PROTEIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/740,036  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0133 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 234 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 544241  
; US-08-740-036-5

Query Match 0.7%; Score 8; DB 4; Length 234;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160  
Db 4 LLLALLV 11

RESULT 40  
US-08-118-270-28  
; Sequence 28, Application US/08118270  
; Patent No. 5508384  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Randall B.  
; APPLICANT: Schuster, David I.  
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF  
; NUMBER OF SEQUENCES: 348  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300

CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/118,270  
FILING DATE: 09-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/943,236  
FILING DATE: 10-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Townsend, Kevin G.  
REGISTRATION NUMBER: 34,033  
REFERENCE/DOCKET NUMBER: MURPHY-2A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 315 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-118-270-28

Query Match 0.7%; Score 8; DB 1; Length 315;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLV 1161  
Db 48 LLLALLV 55

RESULT 41  
PCT-US93-08528-28  
; Sequence 28, Application PC/TUS9308528  
; GENERAL INFORMATION:  
; APPLICANT: New York University  
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF  
; NUMBER OF SEQUENCES: 348  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/08528  
FILING DATE: 09-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/943,236  
FILING DATE: 10-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Townsend, Kevin G.  
REGISTRATION NUMBER: 34,033  
REFERENCE/DOCKET NUMBER: MURPHY-2 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197

TELEPAX: 202-737-3528  
 TELEX: 248633  
 INFORMATION FOR SEQ ID NO: 28:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 315 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 PCT-US93-08528-28

Query Match 0.7%; Score 8; DB 5; Length 315;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1154 LLLALLVL 1161  
 DB 48 LLLALLVL 55

## RESULT 42

US-09-489-039A-7417  
 Sequence 7417, Application US/09489039A  
 Patent No. 6610836

## GENERAL INFORMATION:

APPLICANT: Gary Breton et. al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 FILE REFERENCE: 2709-2004001  
 CURRENT FILING DATE: 2000-01-27  
 PRIOR APPLICATION NUMBER: US 60/117,747  
 PRIOR FILING DATE: 1999-01-29  
 NUMBER OF SEQ ID NOS: 14342  
 SEQ ID NO 7417

LENGTH: 345

TYPE: BPT

ORGANISM: Klebsiella pneumoniae  
 US-09-489-039A-7417

Query Match 0.7%; Score 8; DB 4; Length 345;  
 Best Local Similarity 100.0%; Pred. No. 44;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1151 GGLLLALL 1158  
 DB 76 GGLLLALL 83

## RESULT 43

US-09-046-736-4  
 Sequence 4, Application US/09046736  
 Patent No. 6090582

## GENERAL INFORMATION:

APPLICANT: KIKLY, KRISTINE  
 APPLICANT: ERICKSON-MILLER, CONNIE  
 TITLE OF INVENTION: Sialoadhesin Family Member-3  
 NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ratner & Prestia

STREET: P.O. Box 980

CITY: Valley Forge

STATE: PA

COUNTRY: USA

ZIP: 19482

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/046,736

FILING DATE: 24-MAR-1998

CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/041,885  
 FILING DATE: 02-APR-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Prestia, Paul F  
 REGISTRATION NUMBER: 23,031  
 REFERENCE/DOCKET NUMBER: GH-50019  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 610-407-0700  
 TELEFAX: 610-407-0701  
 TELEX:

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 374 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-046-736-4

Query Match 0.7%; Score 8; DB 3; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 379 GVLGAVG 386

DB 257 GVLGAVG 264

## RESULT 44

US-07-626-618A-17  
 Sequence 17, Application US/07626618A  
 Patent No. 5422265

## GENERAL INFORMATION:

APPLICANT: Van Tol, Hubert H.M.

APPLICANT: Civelli, Olivier

TITLE OF INVENTION: A No. 5422265el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/07/626,618A

APPLICATION NUMBER: US/07/626,618A

FILING DATE: 7 DEC 1990

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 5422265nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 810-221-8317

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 387 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

US-07-626-618A-17

Query Match

0.7%; Score 8; DB 1; Length 387;

Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 45

US-08-086-439C-3  
Sequence 3, Application US/08086439C  
Patent No. 5468615  
GENERAL INFORMATION:  
APPLICANT: Chio, Christopher L.  
APPLICANT: Huff, Rita M.  
TITLE OF INVENTION: A Synthetic Gene for D4 Dopamine  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Upjohn Company, Corp. Intellectual  
ADDRESSEE: Property Law  
STREET: 301 Henrietta Street  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/086.439C  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Darnley Jr., James D.  
REGISTRATION NUMBER: 33,673  
REFERENCE/DOCKET NUMBER: 4700  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 616-385-5210  
TELEFAX: 616-385-6897  
TELEX: 224401  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-086-439C-3

Query Match 0.7%; Score 8; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 46

US-08-056-051-2  
Sequence 2, Application US/08056051  
Patent No. 5516683  
GENERAL INFORMATION:  
APPLICANT: Grandy, David K  
APPLICANT: Bunzow, James R  
APPLICANT: Civeilli, Olivier  
APPLICANT: Van Tol, Hubert H.-M.  
TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/056.051  
FILING DATE: 19930429  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5516683nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092-C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 910-221-5317  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-08-056-051-2

Query Match 0.7%; Score 8; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 47

US-07-928-611-18  
Sequence 18, Application US/07928611  
Patent No. 5569601  
GENERAL INFORMATION:  
APPLICANT: Van Tol, Hubert H.M.  
APPLICANT: Civeilli, Olivier  
TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Allegretti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/928.611  
FILING DATE: 19920810  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5569601nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 810-221-8317

INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-928-611-18

Query Match 0.7%; Score 8; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 48  
US-08-333-977-17  
Sequence 17, Application US/08333977  
Patent No. 5594108  
GENERAL INFORMATION:  
APPLICANT: Van Tol, Hubert H.M.  
APPLICANT: Civeilli, Olivier  
TITLE OF INVENTION: A No. 5594108el Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Allegretti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/333,977  
FILING DATE: 03-NOV-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/626,618  
FILING DATE: 7 DEC 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5594108nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 810-221-8317

INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-333-977-17

Query Match 0.7%; Score 8; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 49  
US-08-449-153-1

Sequence 1, Application US/08449153  
Patent No. 5686255  
GENERAL INFORMATION:  
APPLICANT: Deth, Richard C.  
TITLE OF INVENTION: Compositions and Methods for Diagnosing  
TITLE OF INVENTION: Schizophrenia  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,153  
FILING DATE: 24-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Cartoll, Alice O.  
REGISTRATION NUMBER: 33,542  
REFERENCE/DOCKET NUMBER: RCD95-02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-449-153-1

Query Match 0.7%; Score 8; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 50  
US-08-434-877-3  
Sequence 3, Application US/08434877  
Patent No. 5721132  
GENERAL INFORMATION:  
APPLICANT: Chio, Christopher L.  
APPLICANT: Huff, Rita M.  
TITLE OF INVENTION: A Synthetic Gene for D4 Dopamine  
TITLE OF INVENTION: Receptors  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Upjohn Company, Corp. Intellectual  
ADDRESSEE: Property Law  
STREET: 301 Henrietta Street  
CITY: Kalamazoo  
STATE: Michigan  
COUNTRY: USA  
ZIP: 49001

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette (DS,HD)  
COMPUTER: Gateway 2000, P5-90  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/434,877

QY 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

QY 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 51  
US-08-449-651-1  
; FILING DATE: 1 July 1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Darnley Jr., James D.  
; REGISTRATION NUMBER: 33,542  
; REFERENCE/DOCKET NUMBER: RCD95-02  
; TELEPHONE: 616-385-5210  
; TELEFAX: 616-385-6897  
; TELEX: 224401  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 387 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-434-877-3

Query Match 0.7%; Score 8; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 52  
US-08-487-811A-18  
; Sequence 18, Application US/08487811A  
; Patent No. 5883226  
; GENERAL INFORMATION:  
; APPLICANT: Civelli, Olivier  
; APPLICANT: Van Tol, Hubert H.M.  
; TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,811A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5883226nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 90,1092-L  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-0002  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 387 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-487-811A-18

Query Match 0.7%; Score 8; DB 2; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 53  
US-08-833-703-1  
; Sequence 1, Application US/08833703  
; Patent No. 6080549  
; GENERAL INFORMATION:  
; APPLICANT: DETH, RICHARD  
; TITLE OF INVENTION: METHODS AND MATERIALS FOR THE DIAGNOSIS AND  
; TREATMENT OF SCHIZOPHRENIA AND RELATED DISORDERS  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
; STREET: Ten Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette

QY 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88

RESULT 53  
US-08-833-703-1  
; Sequence 1, Application US/08833703  
; Patent No. 6080549  
; GENERAL INFORMATION:  
; APPLICANT: DETH, RICHARD  
; TITLE OF INVENTION: METHODS AND MATERIALS FOR THE DIAGNOSIS AND  
; TREATMENT OF SCHIZOPHRENIA AND RELATED DISORDERS  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
; STREET: Ten Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/833,703  
FILING DATE: 08-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Heine, Ph.D., Holliday C  
REGISTRATION NUMBER: 34,346  
REFERENCE/DOCKET NUMBER: NU-431XX  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542 2290  
TELEFAX: 617-451 0313  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-833-703-1

Query Match 0.7%; Score 8; DB 3; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 54  
US-08-475-742-4  
Sequence 4, Application US/08475742  
Patent No. 6121015  
GENERAL INFORMATION:  
APPLICANT: O'Malley, Karen L  
APPLICANT: Todd, Richard D  
TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4 Receptor  
FILE REFERENCE: WU 102 CON DIV  
CURRENT APPLICATION NUMBER: US/08/475,742  
CURRENT FILING DATE: 1995-06-07  
EARLIER APPLICATION NUMBER: US 08/261,293  
EARLIER FILING DATE: 1994-06-16  
EARLIER APPLICATION NUMBER: US 08/014,013  
EARLIER FILING DATE: 1993-01-28  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 387  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(387)  
OTHER INFORMATION: Human D4 Receptor Protein  
US-08-475-742-4

Query Match 0.7%; Score 8; DB 3; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 55  
US-09-060-694-18

Sequence 18, Application US/09060694  
Patent No. 6203998  
GENERAL INFORMATION:  
APPLICANT: Civelli, Olivier  
APPLICANT: Van Tol, Hubert H.M.  
TITLE OF INVENTION: A No. 6203998el Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
STREET: 300 South Wacker Drive  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/060,694  
FILING DATE: 15-APR-1998  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6203998nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092-MM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
TELEX:  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-060-694-18

Query Match 0.7%; Score 8; DB 3; Length 387;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 56  
US-09-378-074-18  
Sequence 18, Application US/09378074  
Patent No. 6437114  
GENERAL INFORMATION:  
APPLICANT: Van Tol, Hubert H.M.  
Civelli, Olivier  
TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Allegretti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/378,074  
FILING DATE: 20-AUG-1999  
CLASSIFICATION: <Unknown>

```
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/928,611
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: No. 6437114nan, Kevin E
;; REGISTRATION NUMBER: 35,303
;; REFERENCE/DOCKET NUMBER: 90,1092-B
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312-715-1000
;; TELEFAX: 312-715-1234
;; TELEX: 810-221-8317
;; INFORMATION FOR SEQ ID NO: 18:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 387 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-378-074-18

Query Match          0.7%; Score 8; DB 4; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 57
US-08-261-293-4
; Sequence 4, Application US/08261293
; Patent No. 6486310
; GENERAL INFORMATION:
; APPLICANT: O'Malley, Karen L.
; APPLICANT: Todd, Richard D.
; TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261,293
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,013
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU 102
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6524
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
```

```
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: N-terminal
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: Protein
;; LOCATION: 1..387
;; OTHER INFORMATION: /note= "Human D4 Receptor Protein"
US-08-261-293-4

Query Match          0.7%; Score 8; DB 4; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 58
PCT-US93-07370-18
; Sequence 18, Application PC/TUS9307370
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07370
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US93-07370-18

Query Match          0.7%; Score 8; DB 5; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 59
US-08-056-051-4
; Sequence 4, Application US/08056051
; Patent No. 5516683
; GENERAL INFORMATION:
; APPLICANT: Grandy, David K
; APPLICANT: Bunzow, James R
; APPLICANT: Civelli, Olivier
; APPLICANT: Van Tol, Hubert H.-M.
; TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```



CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/056,051  
FILING DATE: 19930429  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: NO. 5516680nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092-C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 910-221-5317  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 419 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-056-051-4

Query Match 0.7%; Score 8; DB 1; Length 419;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 60  
US-07-928-611-20  
Sequence 20, Application US/07928611  
Patent No. 559601  
GENERAL INFORMATION:  
APPLICANT: Van Tol, Hubert H.M.  
TITLE OF INVENTION: A No. 559601el Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Allegretti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/928,611  
FILING DATE: 19920810  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: NO. 559601nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 810-221-8317  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 419 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-928-611-20

Query Match 0.7%; Score 8; DB 1; Length 419;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88  
RESULT 61  
US-08-487-811A-20  
Sequence 20, Application US/08487811A  
Patent No. 5883226  
GENERAL INFORMATION:  
APPLICANT: Civelli, Olivier  
TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McDonnell Boehrner Hulbert & Berghoff  
STREET: 300 South Wacker Drive  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,811A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5883226nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092-L  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
TELEX:  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 419 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-487-811A-20

Query Match 0.7%; Score 8; DB 2; Length 419;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 62  
US-09-060-694-20  
Sequence 20, Application US/09060694  
Patent No. 620398  
GENERAL INFORMATION:  
APPLICANT: Civelli, Olivier  
TITLE OF INVENTION: A No. 620398el Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McDonnell Boehrner Hulbert & Berghoff  
STREET: 300 South Wacker Drive  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/060,694  
FILING DATE: 15-APR-1998  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6203998nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092-MM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
TELEX:

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:  
LENGTH: 419 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-060-694-20

Query Match 0.7%; Score 8; DB 3; Length 419;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161  
|||  
Db 81 LLLALLVL 88

## RESULT 63

US-09-378-074-20  
Sequence 20, Application US/09378074  
Patent No. 6437114  
GENERAL INFORMATION:  
APPLICANT: Van Tol, Hubert H.M.  
Civelli, Olivier

TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Allegretti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/378,074

FILING DATE: 20-AUG-1999

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/928,611

FILING DATE: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: No. 6437114nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-B

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 810-221-8317

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:  
LENGTH: 419 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-378-074-20

Query Match 0.7%; Score 8; DB 4; Length 419;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161  
|||  
Db 81 LLLALLVL 88

## RESULT 64

PCT-US93-07370-20  
Sequence 20, Application PC/TUS9307370  
GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 22

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/07370

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 419 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US93-07370-20

Query Match 0.7%; Score 8; DB 5; Length 419;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161  
|||  
Db 81 LLLALLVL 88

## RESULT 65

US-09-252-991A-17394  
Sequence 17394, Application US/09252991A  
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONA

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 17394

LENGTH: 453

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-17394

Query Match 0.7%; Score 8; DB 4; Length 453;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1150 LGGLLLLA 1157  
|||

```

Db      426 LGGLLLLL 433

RESULT 66
US-08-056-051-6
; Sequence 6, Application US/08056051
; Patent No. 5516683
; GENERAL INFORMATION:
; APPLICANT: Grandy, David K
; APPLICANT: Bunzow, James R
; APPLICANT: Civelli, Olivier
; APPLICANT: Van Tol, Hubert H.-M.
; TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/056,051
; FILING DATE: 19930429
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5516683nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1234
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-056-051-6

Query Match      0.7%; Score 8; DB 1; Length 467;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1154 LLLALLVL 1161
          |||||
          81 LLLALLVL 88

RESULT 67
US-07-928-611-22
; Sequence 22, Application US/07928611
; Patent No. 5569601
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/928,611
; FILING DATE: 19920810
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5569601nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-928-611-22

Query Match      0.7%; Score 8; DB 1; Length 467;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1154 LLLALLVL 1161
          |||||
          81 LLLALLVL 88

RESULT 68
US-08-487-811A-22
; Sequence 22, Application US/08487811A
; Patent No. 5883226
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,811A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5883226nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-487-811A-22

```

```

Query Match      0.7%; Score 8; DB 2; Length 467;
Best Local Similarity 100.0%; Pred. No. 57;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1154 LLLALLVL 1161
      11111111
Db      81 LLLALLVL 88

RESULT 69
US-09-046-736-2
; Sequence 2, Application US/09046736
; Patent No. 6090582
; GENERAL INFORMATION:
; APPLICANT: KIKLY, KRISTINE
; APPLICANT: ERICKSON-MILLER, CONNIE
; TITLE OF INVENTION: Sialoadhesin Family Member-3
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P. O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,736
; FILING DATE: 24-MAR-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/041,885
; FILING DATE: 02-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-50019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-046-736-2

Query Match      0.7%; Score 8; DB 3; Length 467;
Best Local Similarity 100.0%; Pred. No. 57;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      379 GVLGAVG 386
      11111111
Db      350 GVLGAVG 357

RESULT 70
US-09-060-694-22
; Sequence 22, Application US/09060694
; Patent No. 6203998
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,694
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6203998nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-MM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-060-694-22

Query Match      0.7%; Score 8; DB 3; Length 467;
Best Local Similarity 100.0%; Pred. No. 57;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1154 LLLALLVL 1161
      11111111
Db      81 LLLALLVL 88

RESULT 71
US-09-378-074-22
; Sequence 22, Application US/09378074
; Patent No. 6437114
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/378,074
; FILING DATE: 20-AUG-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/928,611
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6437114nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-B
; TELECOMMUNICATION INFORMATION:

```

TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 810-221-8317  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-378-074-22

Query Match 0.7%; Score 8; DB 4; Length 467;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 72  
PCT-US93-07370-22  
Sequence 22, Application PC/TUS9307370  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/07370  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US93-07370-22

Query Match 0.7%; Score 8; DB 5; Length 467;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88

RESULT 73  
US-08-485-618-103  
Sequence 103, Application US/08485618  
Patent No. 5728533  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5728533rel Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,618  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32797  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 494 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-485-618-103

Query Match 0.7%; Score 8; DB 1; Length 494;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLALL 1158  
DB 442 GGLLLALL 449

RESULT 74  
US-08-605-672-103  
Sequence 103, Application US/08605672  
Patent No. 5817515  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5817515rel Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/605,672  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:

NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32684  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 494 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-605-672-103

Query Match 0.7%; Score 8; DB 2; Length 494;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLIAL 1158  
Db 442 GGLLLIAL 449

## RESULT 75

US-08-482-293A-103  
Sequence 103, Application US/08482293A  
Patent No. 5831029  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5831029e1 Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,293A  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32684  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 494 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-482-293A-103

Query Match 0.7%; Score 8; DB 2; Length 494;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLIAL 1158  
Db 442 GGLLLIAL 449

Search completed: June 24, 2004, 18:06:46  
Job time : 25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 17:53:06 ; Search time 27 Seconds  
(without alignments)  
4232.427 Million cell updates/sec

Title: US-09-980-403-2  
Perfect score: 6224  
Sequence: 1 MDLPRGLVVALSLWPQFT.....FRSARRRREPGLDTPKYLE 1188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 9619526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2159	34.7	1151	2 A45226	integrin alpha-1 c
2	2149.5	34.5	1180	2 A35854	integrin alpha-1 c
3	1910.5	30.7	1178	2 A41442	VLA-2 protein homo
4	1901.5	30.6	1170	2 I45914	integrin alpha 2 s
5	1863	29.9	1181	2 A33998	integrin alpha-2 c
6	1146.5	18.4	1170	2 S03308	cell surface glyco
7	1135.5	18.2	1163	1 RWU1C	cell surface glyco
8	1104	17.7	1153	2 S00551	leukocyte surface
9	1094	17.6	1153	1 RWU1B	cell surface glyco
10	1073.5	17.2	1163	2 I56126	lymphocyte fuction
11	924.5	14.9	1179	2 A53213	integrin alpha-E c
12	799.5	12.8	1035	2 I58409	integrin alpha-9 c
13	789	12.7	272	2 A55348	integrin alpha-1 c
14	738	11.9	1041	2 T31437	integrin alpha-1 c
15	719.5	11.6	1054	2 JC7294	alphan integrin -
16	693.5	11.1	1039	2 A41131	lymphocyte-Peyer's
17	672.5	10.8	1038	2 S06046	integrin alpha-4 c
18	607.5	9.8	1034	2 A36108	integrin alpha-V c
19	606	9.7	1049	2 A27079	fibronectin recept
20	605	9.7	1091	2 A41543	integrin alpha-6 c
21	604.5	9.7	1053	2 S44250	integrin alpha-5 c
22	604	9.7	1051	2 A40021	integrin alpha-3 c
23	602.5	9.7	1073	2 B36429	integrin alpha-6 c
24	594	9.5	1053	2 I55534	VLA-3 alpha subuni
25	587	9.4	1051	2 A35761	cell surface glyco
26	584.5	9.4	1044	2 T10050	integrin alpha-v c
27	568	9.1	1072	2 A38457	integrin alpha-6 c
28	566.5	9.1	1146	2 S40311	integrin - fruit f
29	563	9.0	1044	2 S16516	integrin alpha-8 c

30	546.5	8.8	1226	2 S44834	F54P2.1 protein -
31	544	8.7	1115	2 T09403	integrin alpha cha
32	539	8.7	1115	2 T09433	integrin alpha cha
33	537	8.6	1048	2 A27421	integrin alpha-5 c
34	536	8.6	1045	2 S60571	integrin alpha v c
35	530.5	8.5	1137	2 UC5950	integrin alpha-7 c
36	523.5	8.4	1135	2 I61186	alpha-7 integrin -
37	515.5	8.3	1139	2 S28277	hypothetical prote
38	510	8.2	1037	2 A60163	glycoprotein IIB -
39	502.5	8.1	1106	2 S38783	integrin alpha cha
40	478.5	7.7	1039	2 A34269	integrin alpha-2b
41	470	7.6	1394	2 A29637	position-specific
42	429	6.9	1086	2 T18523	integrin alpha cha
43	420	6.7	191	2 I47230	VLA-2 protein - pi
44	349.5	5.6	764	2 I36916	glycoprotein IIB -
45	260.5	4.2	3124	2 A40020	collagen alpha 1(X

## ALIGNMENTS

### RESULT 1

A45226  
integrin alpha-1 chain - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 15-Sep-2003  
R:Accession: A45226  
R:Breresewitz, R.; Epstein, M.R.; Marcantonio, E.B.  
J. Biol. Chem. 268, 2989-2996, 1993  
A:Title: Expression of native and truncated forms of the human integrin alpha 1 subun  
A:Reference number: A45226; MUID:93155124; PMID:8428973  
A:Accession: A45226  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-1151 <BR1>  
A:Experimental source: hepatoblastoma cell line HepG2  
A>Note: sequence extracted from NCBI backbone (NCBI:P:124326)  
P:142-317/Domain: von Willebrand factor type A repeat homology <YMAL>

Query Match 34.7%; Score 2159; DB 2; Length 1151;  
Best Local Similarity 38.3%; Pred. No. 8e-136;  
Matches 457; Conservative 235; Mismatches 417; Indels 84; Gaps 20;

QY	23	FNMOTRPRVTPGSRTPFFGVTVQOHDISGNKWLAVGAPLETNGYQKTGDVYKCPVHGN	82
DB	1	FNDVKNSTMTSGPVEDMEFTVQYENEGKWLIGSLVQPKNRTGDVYKCPVGRGE	60
QY	83	---CTKMLG-RVTLNSVSEKDNRLGLSLATPKNSPLACSLWSEHGSSVYTTGCM	138
DB	61	SLPCVKLDLPVNTSIPNVTYKENTFGSTLTATNP-NGGFLACGLPYAYRCGLHLYTTGI	119
QY	139	CSRYNSNFRFKTVPALQRCQTYMDIVIVLDGNSIYPMWVEVQVHFLINI LKKFYIGPQQ	198
DB	120	CSDVSPTFQVNSIAP-VQECSTQLDIVIVLDGNSIYPMWDSVTAFLNDLLKRMIDIGPKQ	178
QY	199	IQGVGVQGVGDVHFFHNDYRSVKDVVEAAASHIEQGGTTRTAPGIEFARSEAP--QK	256
DB	179	TQGVGVQGVGVHFFHNDYRSVKDVVEAAASHIEQGGTTRTAPGIEFARSEAP--QK	238
QY	257	GGRGAKKVMIVITDGGSHSDPDLEKVIQOQSERDNTVRYAVAVLGYNNRNGINPTFLNE	316
DB	239	GARGVAKVMIVITDGGSHDNRKLVQDCEDENIQPFSAILGSLYNGNLSTEFVBE	298
QY	317	IKYIASDPDDKHPNVTDEAALKDIVDALGRIPIFSEGT-NKNETSFGLEMSQTGFSHV	375
DB	299	IKSIASEPTEKHFFNVSDALVITVKTLGERIFALEATADQSAASFEMMSQTGFSAHY	358
QY	376	VEDGVLLGAVGAYDMGAVLAKETSAGKVIPLRESVYLKEFFBELKNHGYLVGYTVSVVSS	435
DB	359	SDQWVLLGAVGAYDMGAVVVMQKASQIIIPNTTFNVSTKKNELASLYLGYTVNSATAS	418
QY	436	RQGRVYVAGAPRFNHTGKVLFTMHNNRSLTIHQAMRQOQIGSVFGSITSDVDGDGVT	495

Db 419 SGDVLVYIAGQPRYNHTQVILYRMDG-NIKILQTLSSGEIQSGYFGSILTTTDDKDSNT 477  
Qy 496 DVLVAGAPMYF-NEGHERGKVVYVELAQRNFVNGTLK-----DSHSYQN--- 539  
Db 478 DILLVAGAPMYNGEKEBQGVYVALNQTFRFYQMSLEPIKQCCSSRQHSNCTTNRKE 537  
Qy 540 ---ARFGSSIASVRDLNODSYNDVWVGCAPLEDNHAGAIYFFHGFGRSILKTPKORTAGE 596  
Db 538 PCGARGTAAVAKDLNLDGFDNDIVIGAPLEDDHGGAVIYHSGKTIKRYAQRIPSGG 597  
Qy 597 LATGLQVFGSGIHGQDLNEDGLIDLAVGALGNVILWSPVQVQINASLHPEPSKINIFH 656  
Db 598 DGKTLKFFGSGIHGEMDLNGDLTDVTIGLGAALFWSRDAVAVVWVWTFNFEKNVNIQK 657  
Qy 657 RDCRSRGRATCLAAFLCFTPIFLAPHFQTTTGVRYNATMDERRVTPRAHLDGDDRET 716  
Db 658 KNCHEGKEVTCINAVCFVKLSKEDTIYADLQVRLDLSRLQISRSFFSGTGERKV 717  
Qy 717 NRAVLSSGQELCERINFHYLDADYVYKVPVTFVSEYSLDDPKGMLDDGWPTTLKSVVP 776  
Db 718 QRNITVTKSB--CTGSGFVMLDKHDPODSVRITLDENLTDPNPENGVLDDSLPNSVHEYIP 775  
Qy 777 FVNGCNEDEHCVPLDLVLDARSDDLPTAMEYQCVLRKPAQDCSAYTILSPDTTFFLIESTRQ 836  
Db 776 FAKDCGNKCKISDL-----SLHVATTEKDLILVRSQND 809  
Qy 837 RVAVEATLENRGENAYSTVLNISQSANLQFASL--IQKEDSDGSIBCVWEERLQKQVCN 894  
Db 810 KENVSLVTKTKDSAYNTRTIHVSPNLVFSGIEAIQKSCSN-----HNITCK 859  
Qy 895 VSPFFFAKAVAFRLDSRESKSIPLHLEIELAAGSDSNRSTKEDNVAPLPHLYE 954  
Db 860 VGYPFLRGWVTEKILFQNTSYLMENVTIYISATSDSEEPETLSDNVNISIPVKYE 919  
Qy 955 ADVLFTRSSLSHYEVKLSN---LERYDIGPPFPSCIFRIQNLGLFPIHGMKMITIP 1010  
Db 920 VGLQFVSSASBYHLSIAANETVPEVINSTEDIGNEINFYLIRKSGSPMPKLSIFP 979  
Qy 1011 IATRSGRNLLKRLDFTDEANTSC-----NIGNSTEYRPTVPEGLDRAPOLNSNSD 1064  
Db 980 NMTSNGFVPL-YPTGLSSSENANCRPHIFEDPFSINSKGVKTTSTDLKRGTLDCNTCK 1038  
Qy 1065 VWSINCIRLVPMQEIINFHLLGNLWLSKALKYKSKIMVNAALQRPQSPHIFREEDP 1124  
Db 1039 FAITCNLTSSDISQVNVSLI--LWKPTFKSYFSSNLNLTIRGEL-RSENASLVLSNNQ 1095  
Qy 1125 SRQIEFTISQOE-DWQVPIWITVSTGLGGLLLALLVLRKLGKFPFSARRR 1176  
Db 1096 KRELAIQISQGLPGRVPLWILLISAPAGLLMLLTLALWKIGFRRPLKX 1148

RESULT 2  
A35854  
Integrin alpha-1 chain precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 23-Oct-1990 #sequence\_revision 13-Sep-1991 #text\_change 15-Sep-2003  
C:Accession: A35854; S11243  
R:Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, P.; Carbonetto, J. Cell Biol. 111, 705-720, 1990  
A:Title: Molecular cloning of the rat integrin alpha-1 subunit: a receptor for laminin A  
A:Reference number: A35854; MUID:90338125; PMID:2380249  
A:Accession: A35854  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1180 <ICN>  
A:Cross-references: GB:X52140; NID:G56493; PIDN:CAA36384.1; PID:G56494  
C:Keywords: cell adhesion; cytoskeleton; transmembrane protein  
F:170-345/Domain: von Willebrand factor type A repeat homology <WA2>

Query Match 34.5%; Score 2149.5; DB 2; Length 1180;  
Best Local Similarity 38.0%; Pred. No. 3.6e-135;  
Matches 463; Conservative 242; Mismatches 405; Indels 107; Gaps 24;

Qy 11 WALSLWPQPTDTEFNDTEKPRVIPGSRAPTQGYTVQOHDISGNKVLVVGAPLETNGYOKT 70  
Db 17 WLLTVILGFCVFNVDVFKNSMSPGVEDMFQYTVQOYENEEGKVLGSLVWGQPKART 76  
Qy 71 GDVYKCPVTHGN--CTKLNLG-RVTLNSVSEKDNMLGLSLATNPKNDSFLACPLWS 126  
Db 77 GDVYKCPVGRBRAMPCKLDLPLVNTSIENVTIENMTFGSLVATNP--NGGFLACGELYA 135  
Qy 127 HECGSSYTTGMCVRNSNFRFSKTVAPALORCQYMDIVIVLDGNSIYPMVVEQHFLL 186  
Db 136 YKCGHLRHTTLCSDVSPFPQVNSFAP-VQECSTQLDVIIVLDGNSIYPMVSVIAFLN 194  
Qy 187 NILKFPYIPGQIQGVVQYGYEDVVHFLANDYRSVKDVVEAASHIEORGSTETRTAFGI 246  
Db 195 DLLKQMDIGPKQTVGIVQYGENVTHFNANKYSTREVLVAANKIGRQGLQITFALGI 254  
Qy 247 EPARSEAP--QKGRGKAKKVMIVITDGHSDSDPLEKVIQOSBRDNVTRVAVAVLGYN 304  
Db 255 DTAKEATFTGARGARGVKVMIVITDGHSDHNYRLKQVIOCEBENQRFISALIGHFN 314  
Qy 305 REGINPFLNEIKYIASDPDDPKHFPNVTDBAALXDIYDALGDRIFSLSEGT-NKNETSG 363  
Db 315 RGNLSTEFVBEIKSIASEPTEKHFNFVSDDELALVTIVKALGERIFALEATADQSAASPE 374  
Qy 364 LMSOTGSSHVVEGVLLGAVDNGAVLKETSAGKVIPLRESYLKEPFEELKNHGA 423  
Db 375 MEMSOTGFSAHYSQDWMLGAVDNGTVVMQKANOQVIPHNTTPTQEPAROMEPPLAS 434  
Qy 424 YLGYTVISVWSRQGRV-YVAGAPRPNHTGKVLFTMNNRSLTIHQAMRGQIQISGYFS 482  
Db 435 YLGYTVNS--ATIPGDVLYIAGQPRYNHTGVVYKMDG-NINILQTLGGQISGYFS 491  
Qy 483 EITSVDIDGQVTDVLLVAGAPMYF-NEGHERGKVVYVELAQRNFVNGTLK 530  
Db 492 VLTITDIDKDSYTDLLVAGAPMYGTEKEEQGVVYVAVNQTFRFYQMSLEPIRQCCSS 551  
Qy 531 LKDSHSYQ-----NARFGSSIASVRDLNODSYNDVWVGCAPLEDNHAGAIYFFHGFGRS 583  
Db 552 LKDSCTKENKPCGARGFTAAVAKDLNLDGFDNDIVIGAPLEDDHGGAVIYHSGKT 611  
Qy 584 ILKTPKORTASBELATGQYFCGSIHQDLNEDGLIDLAVGALGNVILWSPVQVQINA 643  
Db 612 IREAVAQRIPSGGDKTLKFFQSGIHGEMDLNGDLTDVTIGLGAALFWARDVAVVKV 671  
Qy 644 SAHFSPSKINIFHRCKSRGRDATCLAAFLCFTPIFLAPHFQTTTGVRYNATMDERRV 703  
Db 672 TNVFPKNVNIQKNCRVGKETVCINATWCFHVKLSKEDSIYEADLQVRLDLSRLQI 731  
Qy 704 PRAHLDGGRFTNRAVLSSQBELCERINFHVLPTADYVYKVPVTFVSEYSLDDPKHGMPL 763  
Db 732 SRSPFSGTQERKIQRNITVRESE--CIRHSFVMLDKHDPQDSVRVTVLDENLTDPNPENGVL 789  
Qy 764 DDGWPTTLVSVFPWNGCNEDEHCVPLDLVLDARSDDLPTAMEYQCVLRKPAQDCSAYTILS 823  
Db 790 DDALPNSVHEH:PFPAKDCGNKCKISDLTLNVSTT----- 824  
Qy 824 FDTVPIIESTRQVRVAVATLENRGENAYSTVLNISQSANLQFASL--IQKEDSDGSIB 881  
Db 825 -EKSLLIVKSQHKENVSLVTKNGDSAYNTRTVVQHSNPLIFSGIEEIQKSCSN--- 880  
Qy 882 VNEERLQKQVCNVSPFPRAKAVAPRLDSFPSKSIPLHLEIELAAGSDSNRSTKE 941  
Db 881 -----QNITCRVGYPFPLRAGETVTFKLIPOFNTSHLSENAIHLISATSDSEEPLES 933  
Qy 942 DNVAFLRFLKYEADVLFTRSSSLSHYEVKLSNLSERY-----DGIGPPFSCIFRIQNLGL 997  
Db 934 DNEVNISIPVKYEVGLQFQYSSASEHHISVAANETIPEFINSTEDIGNEINVTITRKRGH 993  
Qy 998 FPIHGMKMITIPATRSGRNLLKRLDFTDEANTSCINWGNSTEYRPTFVSE--DLRA 1055  
Db 994 FPMPEQLSISFPLNTADGPVLYPIGWSSSD--NVNC-----RPSLDDPFGINS 1043  
Qy 1056 POLNHSNSDV-----SINCIRLVPMQEIINFHLLGNLWLSKALKYKSKIMVNAALQRPQSPHIFREEDP 1100



Db 1044 KWTITKSEVLKRGITQDCSSTCGVATITCSLLPSDLSQVNVSL--LWKPTFRHSS 1101  
Qy 1101 MKIMVAALQRFHSPFFREDPSQIEFEISKQE-DWQYPMIIVGVSTIGLILALL 1159  
Db 1102 LNLTLRGLKSE-NSSLTSSNRKRELAIQISKGLGRVPLMWILLAFAGLLMLL 1160  
Qy 1160 VLALFKLGFPSARRR 1176  
Db 1161 ILALWKIGFEPKPLKK 1177

RESULT 3  
S44142  
VLA-2 protein homolog - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 15-Sep-2003  
C:Accession: S44142  
R:Edelmat, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Danjanovich, L.; Latze  
submitted to the EMBL Data Library, January 1994  
A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not  
A:Reference number: S44142  
A:Accession: S44142  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1178 <EDE>  
A:Cross-references: EMBL:Z29987; NID:G473098; PIDN:CAA82877.1; PID:G473099  
F:169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 30.7%; Score 1910.5; DB 2; Length 1178;  
Best Local Similarity 36.2%; Pred. No. 3.5e-119; Indels 103; Gaps 31;  
Matches 442; Conservative 238; Mismatches 437;

Qy 6 GLUVAWALSLWPGFTD--TFNMDTRKPRVPGSRTAFPGYVQOHDISGNKWLWVGAPLE 63  
Db 8 GALLQLLKLWQGLNCLAYNGLPGAKIFGSPSESGFGYSVQQLTNPGNMLLVGSPWS 67  
Qy 64 TNGYQKTDGVYKCPV--IHGNTCKNL--GRVTLNVSVRKONKRLGLSLATNPKNSFLA 120  
Db 68 GPENRMGDVYKCPVDLPATCEKLNQNSAISNWTIKTNMSGLTLTRNPGTGGFLT 127  
Qy 121 CSPLWSHEGSSYYTTCMSRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSNTPWVE 180  
Db 128 CGPLWAHQNGYATYATGICSDVSPDFLTSFPAVQACPSLVDVVVWCDNSNTPWEA 187  
Qy 181 VQFLINILKFFIIPGQIQGVQGVGDVVEHFLNDRYSVKDVEAAASHTEORGGTET 240  
Db 188 VKNFLVKFVTGLDIPKTKQVALIQVANEPRITPNLDFETKEDWVQATSETROHGGDLT 247  
Qy 241 RTAFGIEFARSAFOK--GCRGAKKQIMIVITDGHSDSPDLKVKVQOSERNVTRYAVA 298  
Db 248 NTFRAIEFADYAYSGTSGRPGATKVMVVDGSHSGSKLTVLQCNDEILRFGIA 307  
Qy 299 VLGYNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAKDIVDALGRIPSLSGTKN 358  
Db 308 VLGYNLRNALDKNLKEIKAIASPTERYFNFADEALALEKAGTGLGEQIFSIQVQG 367  
Qy 359 ETSPLGEMSGTGFSSHV--EDGVLGAVGDMGAVLKETSAGKVIPLRESYLKBPPE 416  
Db 368 GDNFQEMEAQVFSADYAFQNDILMIGAVGDMGTLVQETSHKPV-----PFK 418  
Qy 417 EL-----KNFGAIVGYTTSVSSRQGRVYVAGAPRENHTGKVLFTMHNNRSLTHQ 469  
Db 419 QAQDQVLQDRNHSSPLGYSV-RAISTEDGVHVFAGAPRANYTGQIVLVSNNQGNVTIQ 477  
Qy 470 AMRQOIGSYFSGSEITSDIDGVTDLVAGAPMYFNE-GRERGKVVYVELQRNFVN 528  
Db 478 SHRGDQIGSYFSGSVLCSVDVQDKTITDVLVGLAPTMYNDLKEBKVILFTTIGILNQH 537  
Qy 529 GTLKSHSYQNARFSGSIASVRLNQDSYNDVVVVGAPLEDNHDAGAIYIPHGFSGILKTP 588  
Db 538 QFLEGPECTGNARFSGSIAIALSDINNDGFDVIGSPVENSGAVLYNQHOGTIRTKY 597

Qy 589 KQKITASELA--TGLQVFGCSIHQDLNEDGLIDLAVGALGNVILWSPRVVAINSLH 646  
Db 598 SOKILSGNAGFRRLHQPFGSLDGYDLNGDSITDVSIGALGVQIQLWSQSIADVAIEAL 657  
Qy 647 FBPSKINIHRDCKRSRDATCAALFLCPTPIFLAPHFOITTVGIRYNATWD----BRRY 702  
Db 658 FTPDKITILLNDAK-----ITLKLCFRAEF--RPAGQNNQVAILFNWILDAGHSRV 708  
Qy 703 TPRAHLDGQDRFTNRVALLSSGQELACERINFHVLDTADYVYKPTVFSVYSLEDDHGM 762  
Db 709 TSGVTPRENSERFLQKMMVNEVQK--CSEHHISIQKPSDVVNPDLVDVLSLENGTSPA 767  
Qy 763 LDDGWPTTLRV-SVFPWNGCNEDEHCVPLDLVARSDLPTAMEYCORVLRKPAQCCSAYT 821  
Db 768 L-EAYSETVKVPSIPFYKECGSGIGICISDILDVQ--QLP----- 804  
Qy 822 LSPDTTTFITESTRQVAVEATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGSI 881  
Db 805 -AICQTSFIVSNQKRLTPSVLLKNRGSAYNTVLAEFSENLFASFSPVD--GTEV 860  
Qy 882 VNEERRLOKQV-CNVSPFPRAKAVAPRLDSEFSKSIPLHLHLELAAGSDSNERDSTK 940  
Db 861 TCEVGSSQKSVTCGVYFALKSEQOVTTFINFDNLQNLQAAINQAFSESQE--TNK 918  
Qy 941 EDNVAFLRFLKYEADVLFTRSSLSHVEKLN---SSLERYDGIQPPFCIFRIQ-NL 995  
Db 919 ADNSVSTIPLLYDAEHLTRSTNINFEYISSDENAPSVIKSVEDIGPKF--IFSLKYTA 976  
Qy 996 GLPPIHGMMKTIPIATRSNRLKLRLDFLTDBA--NTSCNINWNSNTRYETPV----- 1048  
Db 977 GSAPVSMALVTIHIPOYTKENPLLYLTGIGTQDQAGDISC-----TABINPLKLPHTAPS 1031  
Qy 1049 ----EEDLRAPQLNHSNDVVSINCNIR-LVNPQENFHLGNLRLSLKALYKSMKI 1103  
Db 1032 VSPKNENFRHTKEDCRTCSTNITCWLKDLHMAEYFINVTRVWNTPAASTFQT--V 1089  
Qy 1104 MVNAALQRFHSPFFREDPSQIEFEISK-QEDWQVPMIIVGVSTIGLILALLAFLA 1162  
Db 1090 QLTAABEIDTNPOLFVIEENAVTIPLMIMKPTKEAEPVTGVIIGSIAGILLAMTAG 1149  
Qy 1163 LRKLGFPRARRRREPGLD 1182  
Db 1150 LMKLGTFK--ROYKKQGNP 1167

RESULT 4  
I45914  
Integrin alpha 2 subunit - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 15-Sep-2003  
C:Accession: I45914  
R:Kanata, T.; Puzon, W.; Takada, Y.  
J.Biol. Chem. 269, 9659-9663, 1994  
A:Title: Identification of putative ligand binding sites within the I-domain of integ  
A:Reference number: A54402; MUID:94193647; PMID:7511592  
A:Accession: I45914  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1170 <KAM>  
A:Cross-references: GB:I25886; NID:G439695; PIDN:AA859255.1; PID:G439696  
F:161-336/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 30.6%; Score 1901.5; DB 2; Length 1170;  
Best Local Similarity 36.3%; Pred. No. 1.4e-118;  
Matches 437; Conservative 238; Mismatches 437; Indels 91; Gaps 30;

Qy 23 FNMTRKPRVTPGSRTPAFPGYVQOHDISGNKWLWVGAPLENTGYQKTDGVYKCPV--IH 80  
Db 19 YNWGLPKAKIFSGSPSESGFGYAVQOQFNPKGNWLLVGSFPMGPPKRNMGDVKYKCPVDLST 78  
Qy 81 GNCTKLNLAGRYT--LSNVSRKDNRLGLSLATNPKNSFLACPSLWSHEGSSYYTTCMC 139  
Db 79 TTCEKLNLTSTSNVEMTKNMSGLTLTRNVGTGGFLTCGPLWAQCCGSSQYTTTGC 138





[illegible]

```

Db      285 AIGVGIAFNRRNS-----WKELNDIASPSQEHIFKVEDFDALKDIOQLQKEKIFAIBG   338
Qy      355 T-NKNETSPGLEMSQTGFSSHHVDEGVLLGAVGAYDWNAGVLKETSAGVIP--LRESYL   411
Db      339 TETTSSTSSFELEMAOBSFGSAVTPDGPVLGAVGSFTWGGAF-----LYPDNMSPTPI   391
Qy      412 KFPPELKDHGAIVGYTTSVSVRSGRQRYVVYAGAPRFNHGTGKVILFTMNNRSLTHQAM   471
Db      392 NMSQENVDNRDSVLYGYS-TELALHWGVQSGLVIGAPRIYHTGRVIFT-QVSRQRWKAEV   449

```

QY 472 RQQTGSGYSGSITSVDIDGQVTDVLLVGVAPMYFNEGRCKVYVYELRQ--NRPVYNG 529  
Db 450 TGTQGSYFGAGSLCSVDVDTGSDTLVLGAPHYEQTR--GGQVSCVPLFRGWRWCD 508  
QY 530 TLKDSHSYQNRFGSSIASVRDLNDSYNDVVGAPLEDNHAGAIYIFHGFRG--SILKTP 588  
Db 509 VLYGQGWGRFGAALTVLGVNGDKLTDVVGAPGEENRGAVLPHSVLGPSPSH 568  
QY 589 KQRTITASELATGLOVFGCSHQDLNEDGLDLAVGALGNVILWSRPVQINASLHFE 648  
Db 569 SORIASGQSSRLQYFGQALSGQDLTDGLVDLAVGARGQVLLLRTPVLVWGVSMQFI 628  
QY 649 PSKINIFHRDCKRS--GRDATCLAAFLCF-----TPIFLAPHFQTTVGIRYNAMDRR 701  
Db 629 PAEIPRSFECREQVQVSEQLTQSNICLVIDKRSKNLLGSRDLQSS--VTLDLALDGR 685  
QY 702 YTPRAHLDEGDRFTNRAVLLSSGQELCERINHLV---DTADYKVPFTSVEYSELD 757  
Db 686 LSPRAFQTKRSLSRVRLGL--KAHCE--NFNLLPSCVDSVTPTLRNFTLVGKP 742  
QY 758 ----DHGPMLDGWPFTTLRVSPVFWNGCNEDEHCVPLDVLDAERDLPTAMEYCORVLRK 812  
Db 743 LIAFNLRLPMLAADAQRYTASLPFEKNGADHIC-----777  
QY 813 PAQDSAYTSLPDTTTFIESTQRVAVATLENGENAYSTVLNISANLQASLI-- 870  
Db 778 --QDNLGISFSPGLKSLVGNLELNAEVMVWMDGEDSYGTTITFSPAGLSYRYVAG 835  
QY 871 QKEDSDGS--ECVNEERLQKQ-----VCNVSPFPRAKAKAERLDSRFS--KSIPLHLE 924  
Db 836 QKQGLRSLHLCDSAPVGSQSTWSTCRINHLIFRGQAIFLAFEDVSPKAVLGDRL 895  
QY 925 IBLAAGDSNERDSTKEDNVAFLRHLKYEAOLVTRSSLSHYEVKNS-- 975  
Db 896 LTAIVSSNNTPRTSK--TTFQLELFPVKY---AVTVVSSHEQFTKYNFSEBESHEV 950  
QY 976 -LERYDGIQFPSCIFRIONGL--PPIHGMWMTIPIATSGNRLKLRLDPLDEANT 1032  
Db 951 AMHY-----OVNLLGORDLV---SINFWPVVE-----LNQEA-- 981  
QY 1033 SCNIWGN-----STEYRTPVEE---DLRAPQLNHSNDVVSINCLRLVP 1076  
Db 982 ---VMDVEVSHQNPPLRCSSEKJAPPASDFLAHQKPNVLDCSIAGCLRPDC--VP 1035  
QY 1077 N---QENFHLGN---WLRSLKALKYKS---KIMVNAALQOFHSPTFRSEDPSR 1126  
Db 1036 SFSVQBELFTLKGNLSPQWVRQILQKQVSVVAVBITPDTSVYSQLPQGEAPRAQTT 1095  
QY 1127 Q18-PEISQEDQWPIWITVIGSTLGLLLALLVLAELKLGFR 1170  
Db 1096 VLEKXVHN-----PTPLIVSSIGLILLALITAVLYKVGFFK 1134

# RESULT 8

S00551  
leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse  
N/Alternate names: complement-3 receptor alpha chain  
C/Species: Mus musculus (house mouse)  
C/Date: 30-Sep-1989 #sequence\_revision 30-Sep-1991 #text\_change 22-Oct-1999  
C/Accession: S00551; 159078  
R/Pytela, R.  
EMBO J. 7, 1371-1378, 1988  
A/Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the f  
A/Reference number: S00551; MUID:88312584; PMID:3044779  
A/Accession: S00551  
A/Molecule type: DNA  
A/Residues: 1-1153 <P>  
A/Cross-references: EMBL:X07640; NID:G52982; PIDN:CAA30479.1; PID:G52983  
A/Note: the authors translated the codon CAC for residue 569 as Glu  
R/Sastre, L.; Roman, J.M.; Teplow, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts,  
Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986  
A/Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep  
A/Reference number: 159078; MUID:86287312; PMID:2942940

A/Accession: 159078  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 11-44 <RES>  
A/Cross-references: GB:M14293; NID:G198993; PIDN:AAA39484.1; PID:G554193  
C/Genetics:  
A/Gene: Mac-1  
C/Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat h  
C/Keywords: cell adhesion; glycoprotein; transmembrane protein  
F1-16/Domain: signal sequence #status predicted <SIG>  
F17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimen  
F148-318/Domain: von Willebrand factor type A repeat homology <WMA2>  
F1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 17.7%; Score 1104; DB 2; Length 1153;  
Best Local Similarity 28.8%; Pred. No. 3e-65;  
Matches 356; Conservative 202; Mismatches 506; Indels 170; Gaps 47;  
QY 5 RGLVAVALSAMPQFTTFNMDTKPRVPGSRFAFFYTVQQSDHDSGNKMLVVCAPLET 64  
Db 4 KALLVT-ALALCHG---FNLDTEHPMTFOENAKG--FGQNVVQ---LGGTSVVAAPQEA 54  
QY 65 NGYKTDGVKCPVHGNCTKLNLRVTLNSVSRKNNRGLGLSLATPKDNSPLACSP 124  
Db 55 KAVNQTAGLYQCDYSTSRCHPIPL-----QVPEAVNMSLGLSLAVSTVQQQLACGPT 108  
QY 125 WSEHCGSYTYTGMCSRVSNN--FRPSKTVAPALQRC-QTYMDIVIVLDGNSI--YPWVE 180  
Db 109 VHONCKENTYVNGLYLPGSNLLRPPQPPALRECPQESDIFVLIDGSGSINNIDPQK 168  
QY 181 VOHFLINILKFFYIGPGQIQGVVQYGEDVHVEHFLANDYRSVKDVEAASHEORGGTET 240  
Db 169 MKEFVSTVMEQF--KSKTLPFLMQYSDEFRIHFTFNDKRNPPSRSHVSIKQLUNG--RT 225  
QY 241 RTAFGIEPARSEAFQK--GGRKGAKVMIVITDGHSDSP--DLEKVIQOQSRDNVTRYAV 297  
Db 226 KTASGIRKVVRELPHKTAGARENAKLLVITDGEKPGDPLDYKQVPEADRAGVIRYVI 285  
QY 298 AVLGYNRRGINPTFLNELIKYIASDDDDKXFFVNTVEAALKDUDALGDSIFSLGNTK 357  
Db 286 GVGNAFNK---POS--REELDTIASKPAGEHVFDNPEALNTIQNQLEKIFAFEGOT 340  
QY 358 NET--SPGLEMSOTGFSHVVEDVLLGAVGAYMNGAVLKETSAGKVIPLRESYLKBPPE 416  
Db 341 GSTSFFHEMSQEFASITNSGFLASVGSFDWAGGAFLYTSKDKV---TPIINTRV 395  
QY 417 ELKNHAYLGVTVTVSVSSRGRVYVAGAPRNFHTGKVLITMNNRSLTIHQMRGQOI 476  
Db 396 DSDMDAYLGY--ASAVILNRVQSLVGLGAPRYQHIGLVWMP--RENFGTWEPHTSIKSGOI 453  
QY 477 GSYFSEITSDIDGQVTDVLLVGVAPMYFNEGRCKVYVYELRQ--RPVYNGTLKDS 534  
Db 454 GSYFASLCSVDMDADGNTNLILGAPHYKTR--GGQVSCVPLFRGWRWCD 512  
QY 535 HSYQNRFPSSIASVRDLNDSYNDVVGAPLEDNHAGAIYIFHGFR--GSILKTPKQRT 593  
Db 513 QGHFWGRFGAALTVLGVNGDKLTDVAIGAPGEQENQAVYIFYGASIASLSASHSHRI 572  
QY 594 ASELATGLOYEGCSHQDLNEDGLDLAVGALGNVILWSRPVQINASLHFEPSKI- 652  
Db 573 GAHFSPLQYFGQSLSGKDLTMDGLMDLAVGAQHLLLRQAQVPLREATWESPKVA 632  
QY 653 -NIFHRDC-----KRSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNAMDRR 704  
Db 633 RSVF--ACQEQVLNKKDAGEVRVCLRVKNTKDLREGDIQST---VTYDLALDPRVSR 687  
QY 705 RAHLDGEGDRFTNRAVLLSSGQELCERINHLVDTA--DYKVPFTSVEYSL--EOP----- 757  
Db 688 RAPFDETKNN--TRRTQTVFGMLQKQCELTGLTLPQCVDDSDVSPILRLNLTVLVGPRLRFG 746  
QY 758 DHGPMLDGWPFTTLRVSPVFWNGCNEDEHCVPLDVLDAERDLPTAMEYCORVLRKPAQC 817  
Db 747 NLRPVLANDQAQRTTAMFPFEKNGDSICQDDLSI-----782







QY 686 -TTVTGIRYNATNDRRYTPRAHLDEGGDRFTNRAVLSSGOELCERINPHV-LDPADYV 743  
Db 663 GRLLANSLYTLQDGHMRSGFLFPDGSHELSGNTSI--TPDKSLDHFHFPIQDULI 720  
QY 744 KPVTSVEYSLEDDHGMDDG-----WPT--TLRVSVFPMWNCNDEHCVPLDVL 794  
Db 721 SPINVLNLSLEEGTPTDQKRAMQPIRLPSIHVTKEIPEKXCKGDKCEANLTL 780  
QY 795 --ARSOLPTAMEYQCRVLKPAQDCSAYTSLSPFTTFFIIESTQRVAVATLENGENAY 852  
Db 781 SPARSG-----PLRLMSSASL-----AVBWTLSNGEDAY 810  
QY 853 STVLNLSQSANTLOF--ASLIQ-----KEDSDGSIKCVNEERLQOV-CNVSYPEF 900  
Db 811 WVRLODFPRGLSFRKVEMLQPHSRMPVSCBELTEGS-----SULTTKLCNVSSPIF 863  
QY 901 RAKAKVAFR-----LDSEFSKSIPLH--HLBIELAAGSDSNEDSTKEDNVAPLPHL 951  
Db 864 KAGQSVSLQWNTLNSWEDFVNLGTVHCE-----NENSSLOEDNSAATHIPV 914  
QY 952 KYEADVL-----FT-----RSSLSH-YEYKLNS-----LRYDGIQPPF 986  
Db 915 LYPVNLTKQENSTLYISPTPKGPKTQQVQHVYQRIQPSAYDHNMPTLEALVGVPRPH 974  
QY 987 SCIFRIQNLGLPPIHGMMKTIPIATRSNRLKLRLDFLTDEANTSCHNMGNSTEYRPT 1046  
Db 975 S-----EDLIT-----YTWVSQTDPLVT 992  
QY 1047 PVTEDLRAPQLNHSNDVVSINCLIRLVNPNQINPHLGNLWL-RSLKALKYSM--KI 1103  
Db 993 CHSEDLK-PSSBARQPLGVQVQPIVFRWEILQVGTVELSKBIKASSTLSLCSL 1051  
QY 1104 MVNAALQRPSPFIFREEDPSQIEFEISKQEDQWQPIWIIVGSLTGLLILALILVAL 1163  
Db 1052 SVSFPSSKXPH--LVGSKASEQVLVKVDLIHEKEM-LHVYLSGIGGLVLLFLIFAL 1107  
QY 1164 RXLGFP-RSARRRE 1177  
Db 1108 YKVGFFKRLKKEWE 1122

RESULT 11  
A53213  
Integrin alpha-B chain - human  
C;Species: Homo sapiens (man)  
C;Date: 19-Oct-1995 #sequence\_revision 31-Mar-2001 #text\_change 31-Mar-2001  
C;Accession: A53213  
R;Shaw, S.K.; Cepek, K.L.; Murphy, B.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.  
J. Biol. Chem. 269, 6015-6025, 1994  
A;Title: Molecular cloning of the human mucosal lymphocyte integrin alpha (B) subunit. Un  
A;Reference number: A53213; MUID:94164962; PMID:8119947  
A;Accession: A53213  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1179 <SHA>  
A;Cross-references: GB:L25851; NID:G457244; PID:G457245  
C;Genetics:  
A;Gene: GDB:ITGAB  
A;Cross-references: GDB:330801  
A;Map position: 17p13  
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
F;199-371/domain: von Willebrand factor type A repeat homology <VMA3>

Query Match 14.98; Score 924.5; DB 2; Length 1179;  
Best Local Similarity 26.3%; Pred. No. 3.2e-53;  
Matches 338; Conservative 206; Mismatches 482; Indels 259; Gaps 52;

QY 23 FNMTRPRVIPSRTAFFGVYQOQHDISGNK-WLVVGAPLETNGYQKTGDVYKCPVIEG 81  
Db 19 FNVDAVRLPMPKGGAPFVLSLLHQDPSTNQWLLVTSF--RTRKTRGFLRCSLQVD 75  
QY 82 NCTKLNLGRVTLNVSERKQNRGLGLSLATNPKNDFLACPLW---SHECGSYTTGM 138

Db 76 BILCHVEHVPPIPKRRHR-----GVTVVRS--HHGVLCIQVLVRRPHSLSSSE--LTGT 125  
QY 139 CS-----RVNSFRFSK-----TVAPALQRC----- 159  
Db 126 CSLLPDLRPOAQANFPLENLLDPDARVDGDCYSNKEGGEDDVNTARQRALEKEBE 185  
QY 160 -----QYMDIVILDCSNSIYP--WVWQHEFLNLLKFKYIGQIQGVGVQ 205  
Db 186 EKDSEEDDEBEAGTBIAIILDSGSGIDPPDFORAKDFISMMRNFKYKFEPCNFALVQ 245  
QY 206 YGEDVVEHFLNDRSVKDVVEAASHIEQRGTETRTAFGIEFARSEAF--QKGRKGAK 263  
Db 246 YGGVITQTEFOLDSQDWMASLARVQNTQVGSV-TKTASAMQHVLDSIFTSSGSRKAS 304  
QY 264 KMIVITDGSNDSP-DLEKVIQOSERNVTRYAVAVLGYNNRGINPDP-----LNEI 317  
Db 305 KMWVLTGGIFEDPLNLTITVINSFQMGXVERPAIGV-----GEEPKSARTAREL 354  
QY 318 KYIASDPDDKHFNFVNTDEAALKDIVDALGDRIFSLGNTNKNETSGLMSQTFSSHYVE 377  
Db 355 NLIASDPDETHAFKVTNTMALDGLLSKLRNIIISMEGTVDALHY--QLAQIGPSAQILD 412  
QY 378 D-GVLLGAVGAYDWN-GAVLKETSAGKVIPLRESYLKEPPELKNHGAYLGYTTSVSS 435  
Db 413 ERQVLLGAVGAFDMGGGALLYDTRSRGRFLNQTAADAAADAAQY-SYLGYAV-AVLHK 470  
QY 436 RQGRVYVAGAPRFNHTGKVILFTMHNNRSLTIHQAMRQOIGSYFGSEITSVIDIDGVT 495  
Db 471 TCSLSYVAGAPQYKHEGAVPEL-QKEGREAFLPVLEGEOMSGSYFSGSLCPVIDMDGST 529  
QY 496 DVLLGAPMYFNEGRERKQVYVEI--RQNPVYNGTUKDSHYONAPFGSIAVRDLN 553  
Db 530 DFLVAAPFYRHVG--EGRVTVYRLSEQDGSFSLARIILSGHPFTNAPFGFMAAAGDLS 588  
QY 554 QDSYNDVVVGAPLE---DNHA--GAIYIHFGRSGILKTPKQRTASELATGLQVFGCS 607  
Db 589 QDKLTDVAIGAPLEGFGADGASFGSVIYNGHWDGLSASQSRTRASTVAPGLQYFGNS 648  
QY 608 IHGQDLNEDGLIDLAVGALGNNAVILMSRPVQVNASLHFFPSKINIHRDCKRSGRAT 667  
Db 649 MAGFDDISGDGLADITVGTLCQAVVFRSRPVVRLKVSMAFTPSALPIFGNG----- 699  
QY 668 CLAAELCTPIPLAHPFQTTVGIR---YNATMDERRYTTPRAHL-----E 710  
Db 700 VVNVELCFE---ISSVTASSGUREALLNFTLVDVQKRRRLQCSVRSCLGCLREWS 756  
QY 711 QGDRFTNRAVLSSGOELCERINPHVLTADYVKEPVTFSVEYSLEDP----DH-GPMLDD 765  
Db 757 SCSQLCEDILLMPTGELCEB-----DCFSNASVKVSQVQLQTPESQTDHPQPIldr 807  
QY 766 GHPTTLRVSPWNCNDEHCVPLVDARSDLPTAMEYCORVLKPAQDCSAYTSLSPD 825  
Db 808 YTEPAIFQLPYEKACKNKLFCVAELQ-----ATVSOQ 842  
QY 826 TTVFIIESTRQRVAVEATLENRGENAYSTVLNLSQASLQFASLIQKEDSDGSIKCVNEE 885  
Db 843 E---LVVGLTKELTLNININSGEDSYNTSMALNYPRLQLKBM-QKXPSP-NIQCDDPQ 897  
QY 886 RRLQOV--CNVSYPPFR---AKAVAFRLDSEFSKSIPLHLELELAAGSDSNEDSTK 940  
Db 898 PVASVLNMCRIHGVLPKXSSAHVSVWQLE---ENAPNRTADITVTVTNSNRRSLA 953  
QY 941 EDNVAPLRLKLYEA-----DVLFTRS--SSLSHYEVKLNSSLERYDGIQPPFCIFRIG 994  
Db 954 NE-THTLQRFHGFVAVLSKPSIMYNTQGLSHKE----- 988  
QY 995 LGLFPIHG-----MMKTIPIATRSNRLKLRLDFLTDEANTSCHNMG--NSTEYRP 1045  
Db 989 -FLPHVHGNLRGASYQLICVPTKLR-GLQVAAVKLTRLTQASTVCT--WSOERACAYSS 1045  
QY 1046 TPVEEDLRAPQLNHSNDVVSINCLIRLVNPNQI-----NPHLGNLWRLSKALKYKAS 1100  
Db 1046 VOHVEWHSVSCVIAADKENVTVAAEISNDHSEBLLKDVTELQILQIEISFN-----KS 1098



Qy	1101	KKIMWNAALQRPQHSFFIFREEDPDSQIBEISQEDQWQVPTWINGSTGGHLLALLV	1160
Db	1099	LVEGLNA-----ENHRKTIWVFLKDEKYH-SLPPIIKGSGVGLLVIVIL	1143
Qy	1161	LALRKLGFPRARRRPGDPTPK	1185
Db	1144	VILFKCGFFK-RKYQLNLESIRK	1166

RESULT 12

I58409  
integrin alpha-9 chain precursor - human  
C-Species: Homo sapiens (man)  
C-Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 31-Dec-2000  
C-Accession: I58409; A49459  
R-Hibbs, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Takahashi, K.  
Oncogene 9, 611-619, 1994  
A>Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in small cell lung carcinoma  
A-Reference number: I58409; MUID:94119603; PMID:8290272  
A-Accession: I58409  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A-Molecule type: mRNA  
A-Residues: 1-1035 <RES>  
A-Cross-references: GB:D25303; NID:9464180; PIDN:BAA04984.1; PID:G533327  
R-Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytela, R.; Sheppard, D.  
J Cell Biol. 123, 1289-1297, 1993  
A>Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel partner protein for p120cas  
A-Reference number: A49459; MUID:94064789; PMID:8245132  
A-Accession: A49459  
A>Status: preliminary  
A-Molecule type: mRNA  
A-Residues: 30-1035 <PAL>  
A-Cross-references: GB:I24158  
C-Superfamily: integrin alpha-4 chain  
C-Key words: glycoprotein; metal binding; transmembrane protein  
F1-1-27/Domain: signal sequence  
Signal predicted <SIG>

Query Match	12.8%	Score	799.5	DB 2	Length	1035			
Best local similarity	24.2%	Pred.	No. 6e-45						
Matches	307	Conservative	177	Mismatches	418	Indels	369	Caps	49
QY	5	RLVAVAWALSLWPGFTDTFNKMTKRPVPGSRTAFGYTVQOHDISGNKWLWVGAPLET	64						
Db	14	RALLLALVVGATP--AGAYNLDPQRPVHQGPADSPFGYAVLEHFHDNTFWLVGAPKAD	71						
QY	65	NGY----QMTGDVYKCPVTHGN---CTKLNLRGRTLNSVS-----ERKDNRLGLSLA	110						
Db	72	SKYSPSVKSPGAVFKCRV-HTNPDRRCTELDMARGKNRGTS CGKTCREDRDDSEWGVSLA	130						
QY	111	TNPK-DNSFLACPLWSHECGSSYYTT-----GMC SRVNSNFRP-SKTVPAPALORCQTY	162						
Db	131	RQPKADGRVLACNRWK---NITYEADHILPHGFCYLIIPSNLQAKGRTLIPCYE-----	182						
QY	163	NDIVIVLDGNSIYPWVEVQHFLINTLKPYIGQIQVGWVQVQYGEDVWHEHFLNDYRSV	222						
Db	183	-----YKSKY-----	187						
QY	223	KDWVEAASHIEQRGGTETRTAFGIEPARSEAPQKGRKGAKKYMIVITQESHSDSPLEK	282						
Db	188	-----GEEGSG-----	193						
QY	283	VIQQRDNTRYAVAVLGYNNRRGINPEFFLNEIKYIASDPDKHFFWVTDEALKDIV	342						
Db	194	-----	193						
QY	343	DALGDRI FSL EGTNKNETSFGLEMSQTVGSSHVVEGVLLGAVGYDMNGAVL KETSAGK	402						
Db	194	-----CQAGIAGFPTEELVVGAPGSGFWAGTI-----K	222						
QY	403	VIPLRE-SYLKEPPEELKNHG-AYLGYTVTSVSSQGRV-YVAGAPRNRHTGKVLFTM	459						
Db	223	VLMNTVTLKLNDEYNNRRVYLYCAYATAGHFSHPSTIDVVGAPODKGVYIFRA	282						

460	QY	HNNRS---	LTHQAMRGQOIGSYFOSSITSVDIDGCVYTDVLLVGAWPYFNEGRGKVY	516
283	Db	DRRGSTLIIKIPQA--	SGKMGSYFOSSLCVADNLNGDLSD--LLVGAPM--FSEIRDSQVT	338
517	QY	VYELBQNRFPVYNGTLKD	SHSYONASFGSSIASVRDLNODSYNDVVVGAPLEDNHGAIAIY	576
339	Db	VYINRGNGALEBQLALT	GDGAYNAHFPGSIAESDLDNDGFPDVAIGAPKEDPAGAVTI	398
577	QY	FHGFGSILKTPKQBITAS	ELATGUYFGCSTHGQLDNEGLDILAVGL--GNAVILW	634
399	Db	YHGDAGGIVPOYSMKLSG	KINPVLRMFGQSISGGIDMDGNGYPDVTVCAPAFMSDSVLLER	458
635	QY	SPRVVQIYNASLHFPBSK	INIFHRDCKRGRDATTCLAAFLCPTPIFLAHPHFOITTVGIRN	694
459	Db	ASPVITVDSI--FLPGS	INITAPOCHDGOQPVNCLNVTTCFS--FHGKHV--PEEIGLNVY	514
695	QY	ATMD-----ERRYTPRAH--	LDEGDRFTNRVALLSSGOELCERINPHV--LDTADYVXDV	746
515	Db	LMADVAKKKGMPVYFVL	GETWQYTEK--LQTYMEETCEHYVAHVAKRVDVISPI	573
747	QY	TFSEVYSLED-----	PDHGPML--DQGWPTLRVSVFPGNCCNDEHCVPDLVLD	794
574	Db	VPEAAYSLSHEVTGEEER	ELPPLTVLAWKKGOKLAQKNQTVFERNC--RSEDCAADLQLO	632
795	QY	ARSDLPAMECQRVLRPEA	DCSAYTILSPDTTVIIESTRQGVAVEATLENRGENAYST	854
633	Db	GKL--LLSMD-----	-----EKLTYALGAVXNISLNTISINLGGDAYDA	671
855	QY	VLNISQSANLOPASLIQ	KEDSDGSECVNEERRLOQCVNVSYPFFRAKAKVAFRLDSEF	914
672	Db	NVSFNVSRELPIFNWQKEE	M--GISCELLESDFLK--CSVGFPEMRSKSY-----EF	721
915	QY	SKSIFLAHLEIB-----	LAAGSDNBRDSTKYENVAPLRPHLYKEADVLT--RSSSL	965
722	Db	SVIFDTSLSGSEEVLSPI	VAQSNTERSESLHNTILVMLPHEVDTSITGMSPTS	781
966	QY	SHYEVKLASS--LERYD	GIGPPSCIFRIQNLGLFPIHGMMKITTIATSGN--RLU	1020
782	Db	FYGESVDAANPIQLDLE	CHFOPIINILOYNTGPTILPGSSVSGISFENRLSSGGAEMP	841
1021	QY	KLRDLTDBANTSCMIN	WNGNSTEYRTPV-----EEDL-----RRAPQLNHSN	1063
842	Db	HVQEMVWQEGKNC	SFOKN-----PTPCIIPQEQENIFHTPIAFFTKSGRKVLCDKPGI	896
1064	QY	DVWSINCMIRLVPQOE--	INFHLGNLWLSLKLKVKSKMKIMYNAALQRFHSPPIFR	1120
897	Db	SCLTAHCNFSALAKEBS	TIQIYMLLN-----TEILKXKSSSVI-----QFNGRAKVK	944
1121	QY	REDPSRQIJEFTISKOED	QWQVPI-----WIIVGSTLGLLILALLVLALRKL	1166
945	Db	VDPALRVVEIAHGNPEE	VTVPFEALHLEPRCYVVGWIIAISLIVGILIFLLAVLLWCM	1004
1167	QY	GFRSGARRRE	1177	
1005	Db	GFR--REYKE	1013	

## RESULT 13

A55348  
integrin alpha-1 - chicken (fragment)  
C-Species: Gallus gallus (chicken)  
C-Date: 10-Feb-1995 #sequence\_revision 10-Feb-1995 #text\_change 15-Sep-2003  
C-Accession: A55348  
R-Kern, A.; Briesewitz, R.; Bank, I.; Marcantonio, E.E.  
J. Biol. Chem. 269, 22811-22816, 1994  
A-Title: The role of the I domain in ligand binding of the human integrin alpha-1beta-  
A-Reference number: A55348; MUID:94357930; PMID:7521332  
A-Accession: A55348  
A-Status: preliminary  
A-Molecule type: mRNA  
A-Residues: 1-272 <KER>  
A-Cross-references: GB:U10114

F;55-230/Domain: von Willebrand factor type A repeat homology <WNA2>

Query Match 12.7%; Score 789; DB 2; Length 272;  
Best Local Similarity 55.9%; Pred. No. 3.2e-45;  
Matches 151; Conservative 53; Mismatches 62; Indels 4; Gaps 3;  
Qy 108 SLATNPKDNLPLSCPLMSHECGSSVYTTGMCVRVNSNFRFSTVAPALQRCOTYMDIVI 167  
Db 2 TLVTPK-GGFLACGLPIAYKQGLRHVTTGCVNSVSTFETVAVAPSVQECTQIDIVI 60  
Qy 168 VLDGNSIYPWVRVQHPHNLILKFFYIGPQIQVGVVQYGEDVVFHFLNDYRSVKDVE 227  
Db 61 VLDGNSIYPWESVTFALNLLNMDIGPOQTGVIGVQGVVHFLNYLTSTTEVMD 120  
Qy 228 AASHIEQGGTETRTAFGIFPASEAFQK--GEGKAKVMIVITDGHSDSDLEKVIQ 285  
Db 121 ALRIRGGTGTMTALGIDTAREAFTEAHGARGVQKVMIVITDGHSDNTRLOEVID 180  
Qy 286 QSRDNRVRYAVVLGYNRGGINPETFLNEIKYIASDPDDHFFNVTDAAALKDIDAL 345  
Db 181 KCDENIQRFAIALGYSRGNLSTEFVEIEKSIASKPTEKHFNVSDALATVIVEAL 240  
Qy 346 GDRIFSLG-TNKNSTSPGLEMSQTFSSH 374  
Db 241 GERIPALEATTQQAASFEMEMSOAGPSAH 270

RESULT 14

T31437  
integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)  
C;Species: Lytechinus variegatus (variegated urchin)  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jan-2000  
C;Accession: T31437  
R;Hertler, P.L.; McClay, D.R.  
submitted to the EMBL Data Library, May 1998  
A;Description: Alpha SU2, a sea urchin integrin which binds laminin.  
A;Reference number: 221035  
A;Accession: T31437  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1041 <HER>  
A;Cross-references: EMBL:AF067658; NID:G3220240; PID:G3220241; PIDN:AAC23572.1  
A;Experimental source: developmental stage embryo  
C;Function:  
A;Description: binds laminin  
C;Superfamily: integrin alpha-2b chain

Query Match 11.9%; Score 738; DB 2; Length 1041;  
Best Local Similarity 23.7%; Pred. No. 7.8e-41;  
Matches 308; Conservative 168; Mismatches 422; Indels 402; Gaps 51;  
Qy 3 LPRGLVAVALLSLWPGTDTFNDTRKPRVPIGSRITAFPGYTYOQHDISGNKWLVTGAP- 61  
Db 1 MARILLSPVAILLLDSTAGFNDLRAPLXKFGPDGSMFGFSAQHRDGNWDNALVGAPE 60  
Qy 62 LETN-GVQKTGDVYK-CPVTHGN--CTKLNLGRV-TLSNVSEKKNRMLGLSLATNPKD 115  
Db 61 ASTNQSGVYTKGAYVYVCRPLSGPSVCEQIPFDKTNINNV--KSNQWFGATVASGGAN 118  
Qy 116 NSFACASP--LWSHECGSSVYTTGMCVRVNSNFRFSTVAPALQRCOTYMDIVIVLDGSN 173  
Db 119 GQILACAPRLVW----- 130  
Qy 174 SIYPWVEVQHPHNLILKFFYIGPQIQVGVVQYGEDVVFHFLNDYRSVKDVEAASHIE 233  
Db 131 -----VETASVTM 138  
Qy 234 QRGGTETRTAFGIFPASEAFQKGRGAKKVMIVITDGHSDSPDLEKVIQOSRDNVT 293  
Db 139 ROEREPTGTCF----- 149  
Qy 294 RYAVAVLGYNRRGINPETFLNEIKYIASDPDDKHFNVTDAAALKDIDALGDRIFSL 353

RESULT 15

JC7294  
alpaph integrin - sea urchin (Strongylocentrotus purpuratus)  
C;Species: Strongylocentrotus purpuratus (purple urchin)  
C;Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 02-Sep-2000  
C;Accession: JC7294

Db 150 -----VGDSDFTNFVNS-----PCQ 165  
Qy 354 GTNKNETSFG-LEMSQTFSSHVED--GVLLGAVGAYDMGAVLKETSAGKVIDLRSY 410  
Db 166 STNREDYDFDKITHCQAGFSGLIFSDNSALVWAGPSYLLQGIIVQSL-----LNSV 219  
Qy 411 LKEFPEELKNHG-----AYLGTYVTSVVSRRQG-RVTVAGAPRHH-TGKVLFTMHN 462  
Db 220 VQATQD--SNTGYTFDMSYRGYSIALGDFNGDGVQDVVVGTPRAESLMLGLVAIPDQNLN 277  
Qy 463 RSLTHQAMRGQIQSGYFGSEITSVDIDGDGTVTDVLLVLCAPMYFNEGR-----EGKVVV 517  
Db 278 Q---ENQVM-GTQIVAVFGYSVTVDINDTVD--LLVGAPYMDGPAIQWEGAGAVV 332  
Qy 518 YELR-----QNRFFVNGTLKDSHYQNAFPGSSIASVEDLNQDSYNDVVVGAPLEDNH 570  
Db 333 YLQNDVPGASNRSLSTLIGGQI--RSRGLSIASIGSNQDGFNDVAIGAPEGDD 390  
Qy 571 AGAVIFHGFGRGSLIKTPKQRTASELA-TGLOVFGCSHGQDLNEDGLDLAVGA--L 627  
Db 391 AGAVIYHGSANGLASTPAQLTPSTLGHSGITTFPSLQGGQDMKRYPOLLVGASGA 450  
Qy 628 GNAVILMSRPVQVJNASLHFEPSKINIHRDCKRGRDATCLAAFLCTPIFLAHPFOT 687  
Db 451 NTAVLIRTPVVSLDATLNTBFIGINLENKIYELA--DGTWTSFIAMCTFTYTGNYLPD 508  
Qy 688 TVGIRYNATMDERRYTPR--AHLDRGDRFTNRVAVLSSGQELCERINPHVLDOT-ADYVK 744  
Db 509 HIDISVTVTDSGIIANRRAMEVDMDSEITKRELAVSTQ-FCDFLRAVYVNSIEDKLT 567  
Qy 745 PVTFSVEVLEPDHG-----PMLDDGWPTLAVSVFPWNGCNEDEHCVDPDLVDARS 797  
Db 568 PKVTLQVLDLNDSESLRQHPHILPIIDWATWSTQTKVSIQNNC-VNNICIPOLDVTVTP 826  
Qy 798 DLPTAMEYQCVLRKPAQDCSAYTSLFDTTFEIBSTRQVAVATELENRGENAYSTVLN 857  
Db 627 NLPN-----IVGTQOELTLDVSLNRRGSDAFQSSLS 658  
Qy 858 ISQSNLOFASLIQKEDSDGSEICVNEERLQKQVNTSYPPFRKAKVAPRLD-SEPSK 916  
Db 659 VTYPIGLQFVRLERKANMDFVTC-SESDLEIITCDTGNPMV-GRNILEFGLTSTPQV 716  
Qy 917 SIFLHLEIETLAAGSDSNRSDTKEDNVAPLFLHLYKRYADVLFTRSSSLSHYEVKLNSSL 976  
Db 717 SGDKDSIEFPYKAESENSSDPNTLENNELNMTVPVTVDCTL--KLSASYPEIVMYSTQ 773  
Qy 977 ERYDGIQPPFSC-----IFRIQMLGLFPFHGMKMITIPIATSGNRLKL 1022  
Db 774 EDY--VWPPFPAKXNASADIGMVMHLYEVNRTSGSNAGEVSLNQWPKQNEDEGVLYFL 831  
Qy 1023 RDLTDEANTSCNI-----WGNSTEYRPTPVEEDLRRAPQLNH--- 1060  
Db 832 LGIMTEEGVT-CQLTQKANKPSGVKLEPSTAKLSNST---TQVSGRKRREPEVAEALA 886  
Qy 1061 -----SNSDVVINCINIRLVPMQETN-----PHLLGNLWLSLK----- 1094  
Db 887 QTDNVITYCASDSCVLINCTI-----DEINASKSVVRILGRFWERTFOKAVSELTPVQQA 941  
Qy 1095 -----ALKYKMKIMVNAALQRFH-----SPPIFEE--DPSRQTEPEISKQEDMQ 1139  
Db 942 TTASASAAVK---TIPYNIPLRDFSDSTKASTLVTEHLVPPVTPPIA-----W- 988  
Qy 1140 VPIWIVSTLGGHLLALLVLAALRKLGFRRARRRREP 1179  
Db 989 ---WIIVSVLGGIILLIILLGLWKCQFP-----BRKKPG 1021

R: Susan, J. M.; Just, M. L.; Lennarz, W. J.  
Biochem. Biophys. Res. Commun. 272, 929-935, 2000  
A: Title: Cloning and characterization of alphaP integrin in embryos of the sea urchin Sc  
A: Reference number: JC7294  
A: Contents: Embryo  
A: Accession: JC7294  
A: Molecule type: mRNA  
A: Residues: 1-1054 <SUS>  
A: Cross-references: GB: AAD55724  
C: Geneticks:  
A: Gene: sualaphap  
C: Superfamily: integrin alpha-2b chain  
C: Keywords: calcium binding; embryo; glycolysis; glycoprotein; heterodimer; transmembran

Query Match 11.6%; Score 719.5; DB 2; Length 1054;  
Best Local Similarity 23.1%; Pred. No. 1.4e-39;  
Matches 292; Conservative 178; Mismatches 428; Indels 365; Gaps 48;

QY 20 TDTFMDTRKRVIFGSRRTAFPGYTVQCHDHSIGNWLVVVGAPLE-----TNGVQKTGD 72  
DB 19 TVAFNFDLRAPVKFDGPGQSLGFSVAQRDQNTDMLIGAPFAPTTQPGVTNG---GA 74

QY 73 VYKCPVI-----HGMCCKNLGNLAVTLNVSERKDNMLGLSLATNPKDNSFLACSP--LWS 126  
DB 75 VYKCPVPLSGSPCEQVFPDTTGNTEVLID-KSNQWFGATLASSGPDGRILLACAPRLVWL 133

QY 127 HECGSSYTTGMCNRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSIYPMVEVQHFLI 186  
DB 134 Q-----TSTISPT----- 141

QY 187 NILKKEFYIGPGQIQGVWVQYGDVHFEFLNDYRSVKDVEAASHIEORGTEETARGI 246  
DB 142 -----DKREPTGTCP-- 152

QY 247 EPARSEAFQKGRGKAKKVMIVITDGEHSDSPLEKVIQOQSERDNVTRYAVAVLYGNRR 306  
DB 153 ----- 152

QY 307 GINPETFLEIKYIASDPDDKHFFNVDTDEALKDIDVALGDRIPLSEGTNKNETSFGLEM 366  
DB 153 -VGHSDFTNFVNSPCQSTDRDLF-----GFDK-----ITH 182

QY 367 SQTGSSHHVVDG--VLLGAVGAYDNGAVLAKETSAGKVIPLRESYLKEFPPE-LKNEGA 423  
DB 183 CQAGSQAQIPSDNSTLWAGAPGSYVLOQIFAQSL-----TLDVSNTPQAVAFDPS 236

QY 424 YLGVTYTSVSSROG-RVTVAGAPRPH-TGKVILFTMHNRSLT-IHQAMEGQOIGSYP 480  
DB 237 YRGYSIALGDFNGDGLDYVVGTPRGESLRGLVAIF-----DQSLVEIITPVVGEQIVSYP 292

QY 481 GSEITSVDTDGQVTDVLLVAGPMYFN-----EGRERGVVYVELRQNRVYVNGTLKDSH 535  
DB 293 GYSVASVDVNGDGLDD-LLVGAPMTNREPATEKKEAGRVYI--LQNA-----DH 340

QY 536 SYQ-----NARFGSSIASVRDLNQDSYNDVVVVGAPLE-DNHAGAIYIFHGFRGSI 584  
DB 341 SLGAPQMLTGKKIRARFGPITSIGDSNQDGFNDVAIGAPYDGEDNSGVVYIYHGSAGEL 400

QY 585 LKTPQRITASBLA-TGLOYFGCSIHQQLDNLNEDGLIDLAVCA--LGNRVILWSPVYQI 641  
DB 401 RUTESQVLTPSELGSDITTFGFSVDGQDMQNDYDPLVWGAESADRAILVTRTPVVL 460

QY 642 NASLHPEPSKINIFRDCX-RSGRDATCLAAFLCTPIPLAPHFQTTTGWIRYNATMDER 700  
DB 461 EAELETIEPIGINLKNKYELPDGTMVTSFVAVACFI--VTGNHL-PARIGISYTLTVDS 517

QY 701 RYTPRAHL---DEGGDRFTNRAVLSSGOELCERINFVLDLT-ADYVKEPTFSVEYSLED 756  
DB 518 ITSGRALLLEVLSQVTKNRN--LVNGMKFCDPLRAYVYVNTIQDKLTPIAVDLQYELTD 575

QY 757 P-----DHGPMLDGQPTTLVAVSPFMWGCNEDEHCVFDLVLDARSDLPTAMEYCORVL 810  
DB 576 ESILPEYILPIINKEANSQTKQVSIQNNC-VNNICIPICIGITVTNLPN----- 625

QY 811 RKPAQDCSAVTLGTFDTTFFIESTRORVAVEATLENGENAYSTVLNISQSANLQFASLI 870  
DB 626 -----IVIGQAOELTLVVSINNGEDAPQSTLAVYYPEGLQYVRL 666

QY 871 QKEDSDGSIKCVNEERLQOVNVSYPFPRAKAKVAFRLD-SFSPKSIIFLHLEIELAA 929  
DB 667 RRANMDFSVTC-TEDSALRMITCDTGNPLV-GKYNLEFGLTLSTLQVSGDKDNIEFYIVA 724

QY 930 GSDSNERDSTKEDNVAPLRFHLKYEADVLFRSS--SLSHVEVKLNSSLERYD----- 980  
DB 725 GSENNEDPNTLDNNELNVTAVIVDATLKLSSASYPEIVTYRVPEDNIVPEPTKNASEA 784

QY 981 GIGPPFSCIFRIQNLGLFPIHGMKMITIPIATRSNRLKLKRLDFTDEANTSCINWNS 1040  
DB 785 DIGMEVWVHLYEVRNTGSSNAAEVTLNIRWPKDEMGDYLFYLLGIMTDEGVT-CQI--SQ 841

QY 1041 TEYRPTVEBDLRAPOLNHSNDV-----VSINCNI- 1072

DB 842 QGANPLGVKLEASTKEQLSNSSTTOVSGRRKREGEYAEALAAQAEPICTPESCVLINCTID 901

QY 1073 RLVPNQBINFHLGNLNLRL-KALKYKSKMKIMWNA--LQRFHSPFIFREE 1122

DB 902 EIKATKSKVVRILGRFWERTFQKAVSEAVPIQVTLASTATATVRSIPYNIPLPMBPTDS 961

QY 1123 DPSQIEFEISKQEDWQVPI-----WIIVSTLGLLILLALLVLALRKLGPFRSARRREP 1178

DB 962 TKASTL---ITABELVLPVSVIAMIIVSVLGGIILLIILGLWKGFF-----ERKKP 1014

QY 1179 GLD 1181  
DB 1015 GED 1017

Search completed: June 24, 2004, 17:57:27  
Job time : 33 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	6188.5	99.4	1189	1	ITAH HUMAN	Q9u4xx5 homo sapien
2	2403.5	38.6	1167	1	ITAG HUMAN	O75578 homo sapien
3	2159	34.7	1151	1	ITAI HUMAN	P56199 homo sapien
4	2149.5	34.5	1180	1	ITAI RAT	P18614 rattus norv
5	1510.5	30.7	1178	1	ITAE MOUSE	Q62469 mus musculus
6	1501.5	30.6	1170	1	ITAE BOVIN	P53710 bos taurus
7	1863	29.9	1181	1	ITAE HUMAN	P17301 homo sapien
8	1146.5	18.4	1161	1	ITAD HUMAN	P20701 homo sapien
9	1139	18.3	1162	1	ITAD HUMAN	Q13349 homo sapien
10	1132.5	18.2	1163	1	ITAH HUMAN	P20702 homo sapien
11	1104	17.7	1153	1	ITAM MOUSE	P05555 mus musculus
12	1093.5	17.6	1152	1	ITAM HUMAN	P11215 homo sapien
13	1073.5	17.2	1163	1	ITAL MOUSE	P24063 mus musculus
14	954	15.3	1167	1	ITAE MOUSE	Q60577 mus musculus
15	924.5	14.9	1179	1	ITAE HUMAN	P38570 homo sapien
16	825	13.3	285	1	ITAI CHICK	Q90615 gallus gall
17	799.5	12.8	1035	1	ITAS HUMAN	Q13797 homo sapien
18	693.5	11.1	1039	1	ITR4 MOUSE	Q00651 mus musculus
19	673.5	10.8	1032	1	ITR4 XENLA	Q91687 xenopus lae
20	672.5	10.8	1038	1	ITR4 HUMAN	P13612 homo sapien
21	611.5	9.8	1130	1	ITAE HUMAN	P23229 homo sapien
22	607.5	9.8	1034	1	ITAV CHICK	P26008 gallus gall
23	605	9.7	1049	1	ITAS HUMAN	P08648 homo sapien
24	604.5	9.7	1053	1	ITAS MOUSE	P11688 mus musculus
25	597.5	9.6	1066	1	ITR3 HUMAN	P26006 homo sapien
26	594	9.5	1053	1	ITR3 MOUSE	Q62470 mus musculus
27	584.5	9.4	1044	1	ITAV MOUSE	P43406 mus musculus
28	580.5	9.3	1066	1	ITR3 CRISP	P17852 cricetidae
29	579	9.3	1091	1	ITR6 MOUSE	Q61739 mus musculus
30	577.5	9.3	1050	1	ITR5 XENLA	Q06274 xenopus lae
31	568	9.1	1072	1	ITR6 CHICK	P26007 gallus gall
32	566.5	9.1	1146	1	ITAI DROME	Q24247 drosophila
33	553	9.0	1044	1	ITR8 CHICK	P26009 gallus gall

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF109681; AAP01258.1; -  
 EMBL; AF137378; AAD51919.2; -  
 EMBL; AL353064; CAB94392.1; -  
 RSP; F17301; IAON.  
 MIM; 604789; -  
 GO; GO:000305; C:integrin complex; TAS.  
 GO; GO:0004895; F:cell adhesion receptor activity; TAS.  
 GO; GO:000518; F:collagen binding; TAS.  
 GO; GO:0007160; P:cell-matrix adhesion; TAS.  
 GO; GO:0007517; P:muscle development; TAS.  
 InterPro; IPR000413; Integrin\_alpha.  
 InterPro; IPR002035; VWF A.  
 Pfam; PF01839; FG-GAP; 3.  
 PRINTS; PR01185; INTEGRINA.  
 PRINTS; PR00453; VWFADOMAIN.  
 SMART; SM00191; Int\_alpha; 5.  
 SMART; SM00327; VWA; 1.  
 PROSITE; PS00242; INTEGRIN\_ALPHA; FALSE\_NEG.  
 PROSITE; PS0234; VWFA; 1.  
 Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 Signal; Repeat; Calcium; Magnesium; Polymorphism.  
 SIGNAL; 1 22  
 CHAIN; 23 1189  
 DOMAIN; 23 1142  
 TRANSMEM; 1143 1165  
 DOMAIN; 1166 1189  
 REPEAT; 38 94  
 REPEAT; 102 163  
 REPEAT; 167 345  
 REPEAT; 359 420  
 REPEAT; 422 475  
 REPEAT; 477 537  
 REPEAT; 539 598  
 REPEAT; 601 653  
 DOMAIN; 1154 1162  
 DOMAIN; 1174 1177  
 CA\_BIND; 488 496  
 CA\_BIND; 551 559  
 DISULFID; 76 83  
 DISULFID; 121 139  
 DISULFID; 129 159  
 DISULFID; 659 688  
 DISULFID; 674 729  
 DISULFID; 781 787  
 DISULFID; 881 893  
 CARBOHYD; 82 82  
 CARBOHYD; 95 95  
 CARBOHYD; 291 291  
 CARBOHYD; 331 331  
 CARBOHYD; 358 358  
 CARBOHYD; 449 449  
 CARBOHYD; 462 462  
 CARBOHYD; 528 528  
 CARBOHYD; 642 642  
 CARBOHYD; 694 694  
 CARBOHYD; 857 857  
 CARBOHYD; 884 884  
 CARBOHYD; 973 973  
 CARBOHYD; 1032 1032  
 CARBOHYD; 1040 1040  
 VARIANT; 433 433  
 VARIANT; 524 524  
 VARIANT; 524 524  
 /FTid=VAR\_009889.  
 /FTid=VAR\_009890.

FT VARIANT 972 972 L -> P.  
 FT VARIANT 1003 1003 I -> M.  
 FT VARIANT 1030 1030 Missing.  
 FT VARIANT 1094 1094 L -> V.  
 FT SEQUENCE 1189 AA; 133609 MW; 5030308A44CD52 CRC64;  
 Query Match  
 Best Local Similarity 99.4%; Score 6188.5; DB 1; Length 1189;  
 Matches 1184; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
 QY 1 MDLPRGLVVAWALSILWPGFTDTFNMDTRKPRVIPSRTAFPGYTVQOHDISGNKWLVA 60  
 DB 1 MDLPRGLVVAWALSILWPGFTDTFNMDTRKPRVIPSRTAFPGYTVQOHDISGNKWLVA 60  
 QY 61 PLENGYQKTGDVYKCPVIHGNCTKLNLRVTLNVSERKDNMRLGLSLATNPKDNSFLA 120  
 DB 61 PLENGYQKTGDVYKCPVIHGNCTKLNLRVTLNVSERKDNMRLGLSLATNPKDNSFLA 120  
 QY 121 CSPLWSHECGSSYTTGKCSRVSNSNFRPSKTVAPALQRCQTYMDIVIVLDSNSIYPWE 180  
 DB 121 CSPLWSHECGSSYTTGKCSRVSNSNFRPSKTVAPALQRCQTYMDIVIVLDSNSIYPWE 180  
 QY 181 VOHFLINILKFKFYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDWEAASHIEQGGTET 240  
 DB 181 VOHFLINILKFKFYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDWEAASHIEQGGTET 240  
 QY 241 RTAFGIEFARSEAPQKGRGAKKMIIVITDGSHTSPDLKVIQOSERDNVTYAVAVL 300  
 DB 241 RTAFGIEFARSEAPQKGRGAKKMIIVITDGSHTSPDLKVIQOSERDNVTYAVAVL 300  
 QY 301 GYNRGINPETFLNEIKYIASDDPKHFFNVTDAAALKDIDVLDALGDRIFSLGNTKNET 360  
 DB 301 GYNRGINPETFLNEIKYIASDDPKHFFNVTDAAALKDIDVLDALGDRIFSLGNTKNET 360  
 QY 361 SFGLMSQTFSSHVVEDGVLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEPPEELKN 420  
 DB 361 SFGLMSQTFSSHVVEDGVLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEPPEELKN 420  
 QY 421 HGAVLGYTVTVSSRGRVYVAGAPRNFHTKVLFTMNNRSLTTHQMRGQIQISYF 480  
 DB 421 HGAVLGYTVTVSSRGRVYVAGAPRNFHTKVLFTMNNRSLTTHQMRGQIQISYF 480  
 QY 481 GSEITSDIDGDGVTDLVNGAPMYFNEGRGRGVYVYELRQNFVYNGTLKDSHSYQNA 540  
 DB 481 GSEITSDIDGDGVTDLVNGAPMYFNEGRGRGVYVYELRQNFVYNGTLKDSHSYQNA 540  
 QY 541 RFSSIASVRDLNODSYNDVWVGAPLEDNAGALYIFHGFRGSILKTPKQBITASBLATG 600  
 DB 541 RFSSIASVRDLNODSYNDVWVGAPLEDNAGALYIFHGFRGSILKTPKQBITASBLATG 600  
 QY 601 LQYFGCSIHGOLDLNEDELIDLAVGALGNVILWSRPVQINASLHFPSPKINIFHRDCK 660  
 DB 601 LQYFGCSIHGOLDLNEDELIDLAVGALGNVILWSRPVQINASLHFPSPKINIFHRDCK 660  
 QY 661 RSGDATCLAAFLCFTPIFLAPHQTTVIGIRYNATWDERRYPTRAHLDEGGDFTRAV 720  
 DB 661 RSGDATCLAAFLCFTPIFLAPHQTTVIGIRYNATWDERRYPTRAHLDEGGDFTRAV 720  
 QY 721 LLSSGQELCBRINPHVLDADYVKPVTFSVEYSLEDPDHGMPLDDGWPTTLRVSPVFWNG 780  
 DB 721 LLSSGQELCBRINPHVLDADYVKPVTFSVEYSLEDPDHGMPLDDGWPTTLRVSPVFWNG 780  
 QY 781 CNEDEHCVPLVLDARSDDLPTAMEYCORVLRKPAQDCSAVTLSPDTTVTFIESTRQVAV 840  
 DB 781 CNEDEHCVPLVLDARSDDLPTAMEYCORVLRKPAQDCSAVTLSPDTTVTFIESTRQVAV 840  
 QY 841 EATLENGENAYSTVLNISQSANLQPASLIQKSDSGSIQVNEERELQKQCNVSPFF 900  
 DB 841 EATLENGENAYSTVLNISQSANLQPASLIQKSDSGSIQVNEERELQKQCNVSPFF 900

QY 901 RAKAKVAFRLDSEFSKSIPLHLEIELELAAGSDSNRSTKEDNVAPLRPHLKYEADVLF 960  
DB 901 RAKAKVAFRLDSEFSKSIPLHLEIELELAAGSDSNRSTKEDNVAPLRPHLKYEADVLF 960  
QY 961 RSSSLSHYEVKLNLSLRYDGGPSCIFRIQNLGLPIHGMWKITIPATASGNLL 1020  
DB 961 RSSSLSHYEVKLNLSLRYDGGPSCIFRIQNLGLPIHGMWKITIPATASGNLL 1020  
QY 1021 KLRLDTLDE-ANTSNCINMGSTYRPTVEEDLRAPQLNHSNSDVWSINCINIRLVPNQ 1079  
DB 1021 KLRLDTLDEVANTSNCINMGSTYRPTVEEDLRAPQLNHSNSDVWSINCINIRLVPNQ 1080  
QY 1080 INFHLGNLWLSLXALKYKSKIMVNAALQROPHSPFIFREDSRQIBFISKQEDWQ 1139  
DB 1081 INFHLGNLWLSLXALKYKSKIMVNAALQROPHSPFIFREDSRQIBFISKQEDWQ 1140  
QY 1140 VPIWIIIVGSTLGLLALLVLAALKLGFRRSARRRBPGLDPTPKVLE 1188  
DB 1141 VPIWIIIVGSTLGLLALLVLAALKLGFRRSARRRBPGLDPTPKVLE 1189

RESULT 2

ITAG HUMAN STANDARD; PRT; 1167 AA.  
AC 075578; Q9DZ8;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-10 precursor.  
GN ITGA10.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Articular chondrocytes;  
RX MEDLINE=98352078; PubMed=9685391;  
RA Camper L., Hellman U., Lundgren-Aakerlund E.;  
RT "Isolation, cloning, and sequence analysis of the integrin subunit  
RT alpha10, a beta1-associated collagen binding integrin expressed on  
RT chondrocytes.";  
RL J. Biol. Chem. 273:20383-20389 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Endothelial cells, and Heart;  
RX MEDLINE=20169197; PubMed=10702680;  
RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,  
RA Wang S.-X., Langley R., Krissansen G.W.;  
RT "The integrin alpha10 subunit: expression pattern, partial gene  
RT structure, and chromosomal localization.";  
RL Cytogenet. Cell Genet. 87:238-244 (1999).  
CC -!- FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10  
CC ASSOCIATES WITH BETA-1.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Widely expressed with highest expression in  
CC muscle and heart. Found in articular cartilage.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VMPA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VMPA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

DR EMBL; AF074015; AAC31952.1; -.  
DR EMBL; AF112345; AAF21944.1; -.  
DR EMBL; AF172723; AAF61638.1; -.  
DR HSSP; P17301; IAOX.  
DR Genew; HGNC:6135; ITGA10.  
DR MIM; 604042; -.  
DR GO; GO:0008305; C: integrin complex; TAS.  
DR GO; GO:0004895; F: cell adhesion receptor activity; TAS.  
DR GO; GO:0005518; F: collagen binding; TAS.  
DR GO; GO:0007160; P: cell-matrix adhesion; TAS.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PRO1185; INTEGRINA.  
DR PRINTS; PRO0453; VMPADOMAIN.  
DR SMART; SM00191; Int\_alpha; 4.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; FALSE\_NEG.  
DR PROSITE; PS0234; VMPA; 1.  
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Signal; Repeat; Calcium; Magnesium.  
FT SIGNAL 1 22  
FT CHAIN 23 1167  
FT DOMAIN 23 1122  
FT TRANSMEM 1123 1145  
FT DOMAIN 1146 1167  
FT REPEAT 38 97  
FT REPEAT ? ?  
FT DOMAIN 167 350  
FT REPEAT 365 427  
FT REPEAT 428 482  
FT REPEAT 483 545  
FT REPEAT 546 605  
FT REPEAT 608 660  
FT DOMAIN 1134 1140  
FT CA\_BIND 494 502  
FT CA\_BIND 558 566  
FT CA\_BIND 620 628  
FT DISULFID 76 86  
FT DISULFID 666 675  
FT DISULFID 681 736  
FT DISULFID 789 795  
FT CARBOHYD 98 98  
FT CARBOHYD 234 234  
FT CARBOHYD 336 336  
FT CARBOHYD 364 364  
FT CARBOHYD 733 733  
FT CARBOHYD 763 763  
FT CARBOHYD 839 839  
FT CARBOHYD 921 921  
FT CARBOHYD 1011 1011  
FT CARBOHYD 1018 1018  
FT CARBOHYD 1039 1039  
FT CONFLICT 844 844  
FT CONFLICT 909 909  
FT CONFLICT 926 926  
SQ SEQUENCE 1167 AA; 127573 MW; A87D3A1C25C1AE0 CRC64;  
Query Match 38.6%; Score 2403.5; DB 1; Length 1167;  
Best Local Similarity 43.0%; Pred. No. 6.4e-148;  
Matches 514; Conservative 212; Mismatches 425; Indels 45; Gaps 16;  
QY 1 MDLPRGLVWALSLWPGFTDFTFMDTRKRVIPGSRVTPGTYVQGHDSGNLWVGA 60  
DB 1 MELPFVTHLFLPLVFLTGLCSPPNLDHHPRLFPGPPEAEFGYSLQHVGGQRMVLVGA 60  
QY 61 PLETNGYQKGDVYKCPV--IHGNCFKLNLGRVTLNSVSKRMELGLSLATNPKDNS 117  
DB 61 PWDGSPDRGDVYRCVPGGNAPCAKHLGDYQLGNSSHPVWMLGKSLLETGDDG 120  
QY 118 FLACSPWISHECGSSYTTGMCGRVNSNFRFSKTVAPALQRCQYTDIVIVLDSNSIYP 177

121	FMACAPLWGRACGSSVFSGICARVNDASFPQOGSLAPTAQRCTYMDVVIVLDGNSIYP	180
178	KVEQORFLINILKFKVIGPQIOIGVVQVQGEDVWHEFHLNDRSVKDVVEAAASHIEQRGG	237
181	WSEVOIFLRLVCKLFDPEQIOVLVQGESPVHEWSLGRFRTKEEVVRAAKNLSRREG	240
238	TETRTAFGIEFARSEAFQK--GORKAKKVMIVITDGHSDPDLEKVIQOQSERDNVTRY	295
241	RETKTAQAMVACTEGFSQSHGRPRBAALLVVVITDGHSDGDEELPAALKACAGRVTRY	300
296	AVAVLGYNNRRGINPTFLMEIKYIASDDPKDHFNVNTOBAALKDIDALGDRIFISLEGT	355
301	GIAVLGHYLRORDPSSFLREIRTIASDDPDERFFNVNTOBAALTDIDALGDRIFGLEGS	360
356	N-KNEVSPGLEMSCTGFSSHWEDZVLGNAGVADNGAVLKETSAGKVIPLRESVLKEF	414
361	HAENESSFGLEMSQIGFSTRLLKDGILFGMWGAYDMGGSVLMLEGGHRLFPFRMALEDEF	420
415	PEELKHGAYLVTVTSVSSRQGVVYAGAPFNHTGKVIPLTMNRESLTIHQAMRQO	474
421	PPALQNHAAVLYGSVSSMLLRGGRLLFLSGAPFRHRGKVIAFOLKDKGAVRVAQSLQCE	480
475	QIGSYRGSEITSDVIDDGGVTDVLLVACAPMVFN--EGREBCKVVVYEL--QNRPFVYNGTLK	532
481	QIGSYRGSELCUPDTRDGTVDLLVAAPMFLQPKETGRVTVILVGGQSLTLQGTILQ	540
533	DHSYQVNRPGSSIASVEDLNQDSYNDVVVVGAPLEDNHAGAIYIFGFRGSLIKTPKORI	592
541	PEPF--QDARPGFAMGALPOLNQGFADVAVGAPLEDHGOALYLHGTSQGVHRHPAQRI	599
593	TASELATGLQYRCSTHGQOLDNEDGLIDLAVGNVILWRSVPVQVNASLHPBPSKI	652
600	AAASMPHALSYFGRSDVRGLUDGDDLDVAVAGQAAILLSRPVHLTPSPLEVTPOAI	659
653	NIFHRDCKRGRDATCIAELCFPTIFLAPHFOITTVGINRYNATMDERYTTPRAHLDEGG	712
660	SVQORDCRRQGEAVCLTAALCFQVTSRTPGRWDHQFYMRFTASLDEWTAGARAAPDGGG	719
713	DRPTNRVILSSGOELCERINEFVLDTADVVKVTSFVSEYSLDEPDH--GPMLDDGWEFTTL	771
720	QRUSPRKRLUSGVNUTCQELHFHVLDTSDFLRPVALTVPFALONTYTKPGFVLNEGFTSI	779
772	RVSVYPWNGCNEDEHCYPDILVDARSDFLTAMEYCORVLRKPAQDCSAYTSLPDTTVFII	831
780	QKLVFPKDCOPDNECVTDLVLQNMDOI-----RGSRK-----APFVV	817
832	ESTQRVAVZATLENRGENAYSTVLNISQSANLQFASLICKEDSDGSIECWNEERLQXQ	891
818	RGRHKVILVSTTLNKRKNAYNTLSLIFRNLHLASLTQORSPKIVECAPRS--AHAR	875
892	VCMVSYPPFRAKAVAPFLDSFESKSIPLHHEIELEAAGSDSNERDSTKEDNVAPLRFHL	951
876	LCSVGEHPVOTGAQVTFLLFEFESCSLLSQVFGKLTASSDSLERNGTLQENTAQSAYI	935
952	KYRADVLFRSSSLSHYEVKLNLSLERYDGIIGPPFSCI PRIONGLGPPHIGMMKTIPI	1011
936	QYEPHLLFSSSESLHRYEHPYGLP--VGPGEFKTTLVQNLCCVWSGLIISALLPA	993
1012	ATRSGRNLLKLRDLTDEANTSCNIGNSTERYPTPVE--BOLRRAPOLNHSNDVWSINC	1070
994	VAHGNYFLSLSQVITN--NASC-IVOKLTEPPGPPVHPELOQTNLGANSNQCQVVR	1050
1071	NI--RLVPNQELNEFHLGNMLRSILKALKYSKMKIWNAAALQORQFSPFIFREDDPSQIE	1129
1051	HLGQAKGTEVSGVLLRLVHNEFFRAKFLSLTVVSTFELGTEGSLVQLTEASRWSESL	1110
1130	FEISKQBDQVPIWIVGSTLGGLLLLALLVIALVRKLGPRFSAR-----RRREPGLO	1181
1111	LEVQTRPILISLILGSLVIGLLALLVFCULWKLGLFFAHKKIPBEEKREKLE	1166

DEPT. 5 3

## RESULTS 3

FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 504 504 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 720 720 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 752 752 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 855 855 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 880 880 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 887 887 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 911 911 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 938 938 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 946 946 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 980 980 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. . .) (POTENTIAL) .  
 SQ SEQUENCE 1151 AA; 127837 MW; 6B3F31AABF52808 CRC64;

Query Match 34.7%; Score 2159; DB 1; Length 1151;  
 Best Local Similarity 38.3%; Pred. No. 4.7e-132;  
 Matches 457; Conservative 235; Mismatches 417; Indels 84; Gaps 20;

QY 23 FMDTRKPRVIGSRAPFTAFYGYTQOHDISGNKWLTVGAPLETNGYQKTDGVYKCPVHGN 82  
 DB 1 FNVVKNSTSGPVEDVFGYVQYENEGKWLIGSLVPGVQPNRTGVDVYKCPVGRGE 60

QY 83 ---CTKILNG-RVTELSNYSBKNNRGLGLSLATNPKNKSLACPLSLHSGSSYTTGM 138  
 DB 61 SLPCVKLDELWNTSIPNTEYKNNMTFGSTLTVPN-NGCFIACGLPLAYRGCGLHYTTGI 119

QY 139 CSRVNSNFRKTPAPALQRCQYMDIVIVLDGNSIYPPVEVQHFLINILKFKVIGPGQ 198  
 DB 120 CSDVSPPTQVNSIAP-VQECSTQLDIVIVLDGNSIYPPWDSVTAFLNLLKRWMDIGPKQ 178

QY 199 IQGVGVQGVGDVWHEFHNDVRSVKDVVEASHIIEQRCGTTRTAFGIEFARSAP--QK 256  
 DB 179 TQGVGVQGVGVWHEFHNLKYSSTEEVLVAKKIVQRCGRQTMALGTDTRKEAFTEAR 238

QY 257 GRRGAKKVMIVITDGBSHDSDLEKVIQOQSERDNVRYAVAVLGYNRRGINPETLINE 316  
 DB 239 GARRGVKVMIVITDGBSHDHRLLKVIQDCEDENIQRSIALLGSYNRGNLSTKTFVBE 298

QY 317 IKYIASDDKHFNVTDEAALKDIVDALGRIPFSECT-NKNETSFGLEKXSGTGFSSHV 375  
 DB 299 IKSIASEPTKHFNVSDDELAVTIVTKLGERIFALREATAQSAASPEMEXSGTGFSAHY 358

QY 376 VEDGVILGAVGAYDNGAVLXETAGKVIPLRESVLKPEPPELKNHGAIVGYTVSVSS 435  
 DB 359 SODWMLGAVGAYDNGVTVVQKASQIIIPNTIFNVESTKKNPLASLYLGTVNSATAS 418

QY 436 RQGRVYVAGAPRPHGTGKVIPLTMNRSALTHQARQQOIGSYFGSGEITSVDIGDGV 495  
 DB 419 SGDVLVIAGQPRYNTGQVILYRVEDG-NIKILQTLGEGIGSYFGSLITTTDIDKDSNT 477

QY 496 DVLIVGAPMYP-NEGREGKRYVVELRQNRVYNGTLK-----DSHYQN--- 539  
 DB 478 DILAVGAPMYPGTEKEEGKRYVVALNQTRFEYQWLSLEPIKQTCSSRQHNCSCTENKNE 537

QY 540 ---ARFGSIASVADLNDSVNDVVGAPLBDNHAGAIYIPHGRGSLIKTKPKORITASE 596  
 DB 538 PCGARFGTALAAVKDLNLDGNDIVIGAPLEDDHGGAVIYHSGKTKIRKYAQRIPSGG 597

QY 597 LATGLQVPGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQINASLHPFSPKINIFH 656  
 DB 597 LATGLQVPGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQINASLHPFSPKINIFH 656

Db 598 DGKTLKFPQSGIHGEMDLNGDLTDVTIGGLGGAALFWSRDVAVVVKVTNFPNKNVNIQK 657  
 QY 657 RDKRSGRDATCLAAFLCFTPIFIAPHFQTTTIGIRYNATMDERRYTPFRAHLDGEGDRFT 716  
 Db 658 KNCHMEGKETVCINATVCFEVLKSKEDTIYEAQLQVRVTLDSLRQISRSFFSGTQBRKV 717  
 QY 717 NEAVLLSSGOELCERINFHVLDTADYKVPYTFVSLEDDHGHMLDGMPTTLRVSVP 776  
 Db 718 QRNITVAKSB--CTKHSFYMLDKHDFQDSVRITLDFNLDPENGLDSDSLPNSVHYIP 775  
 QY 777 FWNGCNDEHCVPDLVLDARSDLPTAMEYQORVLKPAQDCSAYTLQDFTTTFVFIESTRQ 836  
 Db 776 FAXDOGNKCKISDL-----SLHVAITTEKDLLIVRSQND 809  
 QY 837 RVAYEATLENRGENAYSTVNLISOSANLQFASL--IQKEDSDGSGIECVNEERLQKQVCN 894  
 Db 810 KFNYSLVTKTKDSAYNTRTIVHSPNLVFSGLTAEIQAOKSCSN-----HNITCK 859  
 QY 895 VSYOFFRAKAVAPRLDSBFSKSTFLHLELAAGSDNSNERDSTKDDNVAPLFLHLYKE 954  
 Db 860 VGPFPLRGEMVTPKILFQNTSVLMENVILYLSATSDSEPPETLSDNVVNIISIPVYE 919  
 QY 955 ADVLFTRSSSLSHVEVKLNS---LERYDIGIPFSCIPRIQNLGFPFIHGMKMTIP 1010  
 Db 920 VGLQFYSSAGEYHISIAANETVPEVINSTEDIGNEINIFYLIRKSGSGFPPELKSISFP 979  
 QY 1011 IATFSGNRLKLKRLDFTDEANTSC-----NINGNSTERYPTVEEDLRAPQLNHSND 1064  
 Db 980 NMTSGVPLV-YPTGLSSSNANCRPHIFEDPFSINSKGMKWTSTDLKRGITLDCNTCK 1038  
 QY 1065 WVSINCIRLVPQOINFLHLLGNLRLSKALKYKSMKIMVNAALORQHFSPFIREBDP 1124  
 Db 1039 PATITCNLTSSDISQVNVSLI--LWKPTFKSVFSSNLNLTIRGEL-RSENASVLSSNQ 1095  
 QY 1125 SRQIEFEISKOB-DWQVPIHIVGSTLGLGILLALINLALRLKGLGFPFSARRR 1176  
 Db 1096 KRELAIQISKDGLGPRVPLWLLSAPAGULLMLLALWKIGFFRPLKXK 1148

RESULT 4  
 ITAL RAT STANDARD; PRT; 1180 AA.  
 ID ITAL RAT  
 AC P18614;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1)  
 DE (CD49a).  
 GN ITGA1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90338125; PubMed=2380249;  
 RA Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A.,  
 RA Esch F., Carboneito S., Reichardt L.F.;  
 RT "Molecular cloning of the rat integrin alpha 1-subunit: a receptor  
 for laminin and collagen."  
 RL J. Cell Biol. 111:709-720(1990).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.  
 RX MEDLINE=99313197; PubMed=10386626;  
 RA Nolte M., Pepinsky R.B., Vanyaminov S.Y., Koteliarsky V.,  
 RA Gotwals P.J., Karpusas M.;  
 RT "Crystal structure of the alpha1beta1 integrin I-domain: insights into  
 integrin I-domain function."  
 RL PNAS Lett. 452:379-385(1999).  
 CC -1- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND  
 CC COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-  
 CC E-R IN COLLAGEN.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1



CC ASSOCIATES WITH BETA-1.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -1- SIMILARITY: Contains 1 VWFA domain.  
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X52140; CAA36384.1; -  
 CC PIR; A35854; A35854.  
 CC PDB; 1CK4; 03-MAY-00.  
 CC InterPro; IPR000413; Integrin\_alpha.  
 CC InterPro; IPR002035; VWFA.  
 CC Pfam; PF01839; FG-GAP; 3.  
 CC Pfam; PF00357; integrin\_A; 1.  
 CC Pfam; PF00092; vwa; 1.  
 CC SMART; SM00327; VWFA; 1.  
 CC SMART; SM00327; VWFA; 1.  
 CC PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 CC Signal; Repeat; Calcium; Magnesium; 3D-structure.  
 CC SIGNAL 1  
 CC CHAIN 29 1180  
 CC DOMAIN 29 1142  
 CC TRANSMEM 1143 1165  
 CC DOMAIN 1166 1180  
 CC REPEAT 44 103  
 CC REPEAT 44 103  
 CC REPEAT 175 388  
 CC REPEAT 377 432  
 CC REPEAT 433 484  
 CC REPEAT 485 565  
 CC REPEAT 567 626  
 CC REPEAT 629 681  
 CC CA\_BIND 497 505  
 CC CA\_BIND 579 587  
 CC CA\_BIND 641 649  
 CC SITE 1168 1172  
 CC DISULFID 82 92  
 CC DISULFID 687 696  
 CC DISULFID 702 755  
 CC DISULFID 807 813  
 CC DISULFID 877 885  
 CC DISULFID 1029 1062  
 CC DISULFID 1066 1073  
 CC CARBOHYD 100 100  
 CC CARBOHYD 105 105  
 CC CARBOHYD 112 112  
 CC CARBOHYD 217 217  
 CC CARBOHYD 317 317  
 CC CARBOHYD 341 341  
 CC CARBOHYD 402 402  
 CC CARBOHYD 418 418  
 CC CARBOHYD 459 459  
 CC CARBOHYD 531 531  
 CC CARBOHYD 698 698  
 CC CARBOHYD 747 747  
 CC CARBOHYD 779 779  
 CC CARBOHYD 820 820  
 CC CARBOHYD 839 839  
 CC CARBOHYD 882 882  
 CC CARBOHYD 907 907  
 CC CARBOHYD 938 938  
 CC CARBOHYD 965 965

FT CARBOHYD 973 973 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1007 1007 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1084 1084 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1103 1103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1114 1114 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1180 AA; 130808 MW; 8E5DA2BE02362EE4 CRC64;  
 Query Match 34.5%; Score 2149.5; DB 1; Length 1180;  
 Best Local Similarity 38.0%; Pred. No. 2e-131;  
 Matches 463; Conservative 242; Mismatches 405; Indels 107; Gaps 24;  
 QY 11 WALSLEPGTDTFNNMDTRKPRVIGPSRTAFQYVQOQHDISGNKVLVVGAPLETNGYOKT 70  
 DB 17 WLTWILGFCYFNVVDKNSMSPSPVEDMFGYTVQOYENBEGKVLGSPVLGQPKART 76  
 QY 71 GGVYKCPVTHGN---CTKLNLG-RVTLNVSERKDNMLGLSLATNPKNDSNPLASCPLWS 126  
 DB 77 GGVYKCPVGRERAMPVCXLDLPVNTSIPNVTIEKENMTFGSLVLTNP-NGGFLACGLPLYA 135  
 QY 127 HECGSSYVTTGMCVRVNSNERSFKTVAPALORCQYMDIVIVLDGNSNLYPMVEYQHFLLI 186  
 DB 136 YRCGHLHYTTGICSDVSPFPQVNSFAP-VQECSTQLDIVIVLDGNSNLYPMVESVIAFLN 194  
 QY 187 NILKFFYTGPGQIQGVGVQYGEDVHVHFLNDYRSKVDVVEAAASHIEQGGTETETAFQI 246  
 DB 195 DLLKRMDIGPKQTQVGIQVQGENVTHEFNLNKYSSTBEVLVAANKIGRQGLQMTALGI 254  
 QY 247 BEARSEAF--QKGRKGAKWMIIVITDGHSDHSDPLEKVIQOSERDNTRYAVAVLYGN 304  
 DB 255 DTAKEAFTEARGARRGVKVMVIVDGHSDHNYRLKQVIOQCEDENQRFSAIAGHYN 314  
 QY 305 RRGINPETPLNEIKYIASDPDDKHFFNVITDEALXDIVDALGDRIFSLSGT-NKNETSPG 363  
 DB 315 RGNLSTKFEVEIKSIASEPTKHFNFVSDLELALVTIVKALGERIFALEATADQAASPE 374  
 QY 364 LEMSGTGFSSHVVEDGVLGAVGADWNGAVLKEFSAGKVIPLRSYLKEPPEELKNHCA 423  
 DB 375 MEMSGTGFSAHYSQDVMGLGAVDWMGTVMQKQVVMQKQVVMQKQVVMQKQVVMQKQV 434  
 QY 424 YLGYTVTSVWSRQGRV-YVAGAPFNHTGKVLFTMNNRSLTHQAMRGQOIGSYFGS 482  
 DB 435 YLGYTVNS--ATIPGDVLYIAGPRYNTGQVVIYKMDG-NINILQTLGGGQIGSYFGS 491  
 QY 483 EITSVDIDGQGVTVLLVIGAPMYF-NEGREGKVVVYELRQNRVYN-----GT 530  
 DB 492 VLTTHIDDKSYTDLVVGAPMYGTEKEGQVYVAVNTRFQYQMSLEPIRQTCSS 551  
 QY 531 LKDSHSYO-----NARFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIPHGRGS 583  
 DB 552 LKONSTKENKNEPCGARFGTAIAAKDLNVDGVDVIGAPLEDHAGAVIYHGSCKT 611  
 QY 584 ILKTPKQITASELATGLQYFGCSIHGOLDINEDGLIDLAVGALGNAILSRPVVQINA 643  
 DB 612 IREAVAQRIPIPGGDKTKLFFQSIHGEDMLNGDGLTDVTIGLGGALFWARDVAVVKV 671  
 QY 644 SLHPEPSKINIFHRDCKESGRDATCAAPLCTPFIPLAPHQPTTVTVGIRYNATMDERYT 703  
 DB 672 TNHFEFNKVNQKQKCRVEGNETVCINATMCFHVKLSKEDSIYEADLOQYRVLDSLQI 731  
 QY 704 PRAHLEGGDRFTNRAVLLSSQELCERINHFVLDADYVVKPFTSVESYSLDPOHGM 763  
 DB 732 SRSPSGTQERKIQRNITVRESE--CIRHSPYMLDKHDPQDSVRVTLDFNLTDPEGPVL 789  
 QY 764 DCGWPTTLRVSVFPMWNGCNEDEHCVPLDALVARSDLPTAMEYQCRVLKPAQDCSAYLS 823  
 DB 790 DDLNPNVHEHIPFANDCGNKERCISDLTAVSIT----- 824  
 QY 824 FDTTVFIESTQRVAVVATLENRGENAYSTVLNISQSANLQFASL--IQKEDSDGSIK 881  
 DB 825 -EKSLLIVKSQHDKNFVSLTVKNKXGDSAYNTRTVVQHSPLNIFSGIEEIQKSCESN--- 880  
 QY 882 VNEERRLQKQVNCVNSYPPFRKAKAVAFRLDSEFSKIFLHLELELAAGSNERDSTKE 941

Db 881 -----QNTCRVGVFFRAGTGVTKIFQFNTSHLSENAILHLSATSDSEPLESLN 933  
Qy 942 DNVAPLRFHLKYEADVLFTRSSLSHYEVLKASSLRY-----DGIGPPSPSCIFRIQNLGL 997  
Db 934 DNEVNISIPVYEVGVGFQYSSASEHHISVAANETIPEFINSITDIGNINEINVTIRKRGH 993  
Qy 998 FPIHGMXXITPIATRSGRNLLKLDFTDANSCNWNWSTYRTPVPEZ--DLRRA 1055  
Db 994 FPMPEQLSIFPNLTADGVPVLPYIGWSSD--NVNC-----RPSLEDPFQINS 1043  
Qy 1056 POLNEHNSDW-----SINCRLVFNQBEINPHLLCNLWLRSLKALKVKS 1100  
Db 1044 KGMTSKSEVLRKGTIQDCSSCGVATITCSLLPSLSQVNSLL--LKPFFIRAHFSS 1101  
Qy 1101 KXIMMNAALQROFHPSPFIREDDPSQIEFEISKQB-DWQVPIWIVGSLTGLLLALL 1159  
Db 1102 LNLTLRGLKSE--NSSLTLSSNRKRELAIQISKGLGRVPLWILLFAFAGLLMLL 1160  
Qy 1160 VLALRKIGFFRGARRR 1176  
Db 1161 ILALWKIGFFRPLKK 1177  
RESULT 5  
IT2A MOUSE  
ID IT2A MOUSE STANDARD; PRT; 1178 AA.  
AC Q62469; Q62163;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GP1a)  
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).  
GN IT2A2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6 X CBA; TISSUE=Lung;  
RX MEDLINE=94363406; PubMed=8081889;  
RA Edelman J.M., Chan B.M., Uniyal S., Onodera H., Wang D.Z.,  
RA Hamjanovic L., Latzer D.B., Finberg R.W., Bergelson J.M.;  
RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but  
RT not virus binding."  
RL Cell Adhes. Commun. 2:131-143(1994).  
RN [2]  
RP SEQUENCE OF 450-1178 FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=94355691; PubMed=7521231;  
RA Wu J.E., Santoro S.A.;  
RT "Complex patterns of expression suggest extensive roles for the alpha  
RT 2 beta 1 integrin in murine development."  
RL Dev. Dyn. 199:292-314(1994).  
CC -1- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING  
CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO  
CC COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,  
CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN  
CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION  
CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.  
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2  
CC ASSOCIATES WITH BETA-1.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -1- SIMILARITY: Contains 1 VWFA domain.  
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
EMBL; Z29987; CA82877.1; -;  
EMBL; X75427; CA53178.1; -;  
PIR; S44142; S44142.  
HSSP; P17301; IAOX.  
MGI; 96600; Itga2.  
InterPro; IPR000413; Integrin\_alpha.  
InterPro; IPR002035; VWF A.  
Pfam; PF01839; FG-GAP 3.  
Pfam; PF00357; Integrin\_A; 1.  
Pfam; PF00092; vwa; 1.  
SMART; SM00191; Int\_alpha; 4.  
SMART; SM00327; VWF; 1.  
PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
PROSITE; PS0234; VWFA; 1.  
Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
Platelet; Signal; Repeat; Calcium; Magnesium.  
FT SIGNAL 1 26  
FT CHAIN 27 1178  
FT INTERIN ALPHA-2.  
FT DOMAIN 27 1129  
FT TRANSMEM 1130 1151  
FT DOMAIN 1152 1178  
FT REPEAT 42 100  
FT REPEAT ? ?  
FT DOMAIN 185 375  
FT REPEAT ? ?  
FT REPEAT 431 483  
FT REPEAT 485 546  
FT REPEAT 548 607  
FT REPEAT 612 664  
FT CA\_BIND 496 504  
FT CA\_BIND 560 568  
FT CA\_BIND 624 632  
FT SITE 480 482  
FT SITE 1154 1158  
FT DISULFID 80 89  
FT DISULFID 677 734  
FT DISULFID 786 792  
FT DISULFID 862 873  
FT DISULFID 1016 1047  
FT DISULFID 1052 1057  
FT CARBOHYD 102 102  
FT CARBOHYD 109 109  
FT CARBOHYD 429 429  
FT CARBOHYD 457 457  
FT CARBOHYD 472 472  
FT CARBOHYD 696 696  
FT CARBOHYD 1054 1054  
FT CARBOHYD 1071 1071  
FT CARBOHYD 1078 1078  
SQ SEQUENCE 1178 AA; 128925 MW; 1F194B9C0240F465 CRC64;  
Query Match 30.7%; Score 1910.5; DB 1; Length 1178;  
Best Local Similarity 36.2%; Pred. No. 6.6e-116;  
Matches 442; Conservative 238; Mismatches 437; Indels 103; Gaps 31;  
Qy 6 GLVVAWALSLEPGFTD--TFNMDTRKPRVIGSRFAFGYTVQQHDSIGNKWLWVVGAPLE 63  
Db 8 GALLQLMLVQGLINCLAYNVGLPKAIFGSPSEQFGYVQQLTNGQWLLVGSFWS 67  
Qy 64 TNGYQKTGDEVKCPV--THGNTKLN--GRVTLNVSVSRKONMRGLSLATNPKNDSFLA 120  
Db 68 GPPENRMGVYKCPVDLPATCEKLNQNSASISNWTBIKTMSLGLTLTRNPGTGGELT 127  
Qy 121 CSPLWSHECGSYTTGMCNRSVNSFRSKTVAPALQRCQTYMDIVIVLDGNSNLYPWVE 180  
Db 128 CQPLWAHQGQYATGICSDVSPDFQLTSPSPAVQACPSLVDVVVVCDESNLYPWEA 187  
Qy 181 VQHFLLINLKRFGYIGQVGVQYGVQGVVHFLNDYRSVKDYVVEAASHIEQGGT 240

Db 188 VKNPLVAVTGLDIPKPKTQVALIQYANEPRIIFNLNDFETKEDMYQATSETRHQGDLT 247  
Qy 241 RTAFGLFARSEAFQK--GRKGAKKVMIVITGESHDSPDLKVKVCOSEEDNVTYAVA 298  
Db 248 NTRALFADYATSGRGPATKVMVVVITGESHDGSKLTKVITQQCNDLILFGIA 307  
Qy 299 VLGYNRRGINPTEFLNEIKYIASDPDDKHFNFVNTDEAALKOTVDALGORIFSELTNKH 358  
Db 308 VLGVLNENALDTRNLEIKAIKIASTPTERYFFNVADEAALLKXAGTLGEOIFSEGTVOG 367  
Qy 359 ETSFGLSEMSQTGRSSHV--EDGVLLGAVGAYDNGAVNLKETSAGKVIPLRESYLKEFFE 416  
Db 368 GDNFQEMAGVGSADYAPQNDILMLGAVGAFDMSGLVQETSHKPI-----FFK 418  
Qy 417 EL-----KHGAYLGYTVTSVVSQGRGVYVAGAPRFNHTGKVLFTMHNRSITIHQ 469  
Db 419 QAFDVLQDRHSSFLGYSV-AAISTEDGVHFVAGAPRANYTGQIVLYSVNKGQNVTVIQ 477  
Qy 470 AMGQQIGSGVGSSEITVDIDGQVTDVLLVGAPMYFNE--GRERGKVVYVELRQNFVN 528  
Db 478 SHRGDQIGSGVGSVLCSDVDKDTITDVLVAGAPTYNDLKKBEKGKVLFTITKGLNQH 537  
Qy 529 GTLKDSHYQARFGSGIASVRLNODSYNDVVVVGAPLEDHAGAYIHFGRGSLKTP 588  
Db 538 QFLEGPCTGNARFGSAIALSDINDGDFNDVIVSGPVENSGAVIYNGHGTIRTKY 597  
Qy 589 KQITASELA--TGLQVFGSGISGQLDNLDEGLDLAVGALGNANVILMSRPVQINASH 646  
Db 598 SQKILGNGAFRRHLQFGRSLDGYGLNGDSITDVSIGALGQVILMSQSIADVAIEAL 657  
Qy 647 PERSKINIFRDCRGRDATCLAAFCFTPIFLAPHFTTTVIGIRVATMD---ERRY 702  
Db 658 FTEPKITLNDKAK-----ITLKLCFRAEP-RPAGQNNQVAILFNMTLADGHSRV 708  
Qy 703 TPRAHLDEGDRFTNRAVLSSQELCERINFHVLDTADYVVPVTSVEVSLDPDHGPM 762  
Db 709 TSGVFPRENERFLQKNWVNEVQK-CSEHHSIQKPSDVVNPLDLRVDISELNPCTSEA 767  
Qy 763 LDGQWPTLAV-SVPFWNGNEBHCVPDLVLAARSLPTAMEVCQVLEKPAQDCSAYT 821  
Db 768 L-EAYSTVKTFSIFFKESGDSGICISDLILVQ-QLP-----804  
Qy 822 LSPDTTVFIIESTRQVAVATELENRGENAYSTVNLISQSANIOPASLIQKEDSGSIEC 881  
Db 805 -AIQTOSFIVSNQKRLTFSVILKNRGSAYNTVLAARFSENLFASFSMPVD---GTEV 860  
Qy 882 VNEERLQKV-CNVSYPPERAKAVARLDSFESKIFLHLEIELAGSDSNERDSTK 940  
Db 861 TCEVSSQKSVTCDVGPALKSEQQVTFINFDNLQNLQNLQNLQNLQNLQNLQNLQNLQ 918  
Qy 941 EDNVAPRLPHLYEADVLFRSSSLSHYVKLN-----SSLERYDGGIPPFSCIFRIQ-NL 995  
Db 919 ADNSVSLTIPLLYDABLHTRSTNINFEISSDENAPSVIKSVEDIGPKF--IFSLKVTA 976  
Qy 996 GLPFIHGMKKIPIATRSNRLKLRDLFTEA-NTSCNINWGNSTERYPTV-----1048  
Db 977 GSAPVSNALVTIHPQYTKENPLLYLTGTQDQAGDISC-----TABINPLKPTAPS 1031  
Qy 1049 ----EEDLRRAPOLNHSNDSVVSINCNIR-LVPNQEINFLHNLWLSLKLKLYKSMKI 1103  
Db 1032 VSPKNEFRHTKELDCITSCSNITCWLKDLHKAIFYINVTWNRTFAASTPQT--V 1089  
Qy 1104 MVNAAORQSHSPPIFREDDPSQIEFEISK-QEDMQVPTWITVSTGLGLLLALLILVA 1162  
Db 1090 QLTAASIDITHNPOLFVIEENAVTIPLMIMKPEKAEVPTGVITGIIAGILLAMTAG 1149  
Qy 1163 LRKLGFRSARRRREPGLDP 1182  
Db 1150 LKLGFFK--ROYKMGQNP 1167

RESULT 6

ITA2 BOVIN  
ID ITA2 BOVIN STANDARD; PRT: 1170 AA.  
AC P53710;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)  
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).  
GN ITGA2.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
SEQUENCE FROM N.A.  
RP MEDLINE=94193647; PubMed=7511592;  
RX Kamata T., Puzon W., Takada Y.;  
RA "Identification of putative ligand binding sites within I domain of  
RT integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";  
RL J. Biol. Chem. 269:9659-9663(1994).  
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,  
CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND B-CADHERIN. IT  
CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-E-R IN  
CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER  
CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE  
CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
CC EXTRACELLULAR MATRIX.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- ASSOCIATES WITH BETA-1.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L25886; AAB59255.1; -.  
DR PIR; I45914; I45914.  
DR HSSP; P17301; LAOX.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWFA.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; VWFA\_1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS0234; VWFA; 1.  
DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.  
FT NON TER 1 1  
FT SIGNAL <1 18  
FT CHAIN 19 1170  
FT DOMAIN 19 1121  
FT TRANSMEM 1122 1143  
FT DOMAIN 1144 1170  
FT REPEAT 34 92  
FT REPEAT ? ?  
FT DOMAIN 177 367  
FT REPEAT ? ?  
FT REPEAT 423 475  
FT REPEAT 477 538  
FT REPEAT 540 599  
FT REPEAT 604 656

FT	CA BIND	488	496	POTENTIAL.
FT	CA BIND	552	560	POTENTIAL.
FT	CA BIND	616	624	POTENTIAL.
FT	SITE	472	474	CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE	1146	1150	GFEKR MOTIF.
FT	DISULFID	72	81	BY SIMILARITY.
FT	DISULFID	663	726	BY SIMILARITY.
FT	DISULFID	778	784	BY SIMILARITY.
FT	DISULFID	854	865	BY SIMILARITY.
FT	DISULFID	1008	1039	BY SIMILARITY.
FT	DISULFID	1044	1049	BY SIMILARITY.
FT	CARBOHYD	94	94	N-LINKED (GLCNAC. .)
FT	CARBOHYD	101	101	N-LINKED (GLCNAC. .)
FT	CARBOHYD	332	332	N-LINKED (GLCNAC. .)
FT	CARBOHYD	421	421	N-LINKED (GLCNAC. .)
FT	CARBOHYD	449	449	N-LINKED (GLCNAC. .)
FT	CARBOHYD	464	464	N-LINKED (GLCNAC. .)
FT	CARBOHYD	688	688	N-LINKED (GLCNAC. .)
FT	CARBOHYD	748	748	N-LINKED (GLCNAC. .)
FT	CARBOHYD	945	945	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1063	1063	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1070	1070	N-LINKED (GLCNAC. .)
FT	VARIANT	580	580	R -> V.
FT	VARIANT	588	588	G -> K.
FT	VARIANT	725	725	R -> S.
FT	SEQUENCE	1170 AA;	128929 MW;	BECEFLC5F2448FB1 CRC64;
Query Match		30.6%;	Score 1901.5;	DB 1; Length 1170;
Best Local Similarity		36.3%;	Pred. NO. 2.5e-115;	
Matches 437;		Conservative 238;	Mismatches 437;	Indels 91; Gaps 30;
QY	23 FNMTRKRRVPGSRRTAFEGYTVOQHDSGNMKNLVGVGAPLETNGYQYGTGVYKCPV--IH 80			
DB	19 YNVLGPKAKIPSGPSSEDFGVAVQQFINPKGNWLLVSPWSPGFKNRMEDVIKCPVDLST 78			
QY	81 GNCTKLNLGRVT-LSNVSRKDNRLGSLATNPKNMSFLACSPLMWSHECGSSYYTGMCC 139			
DB	79 TTCEKLNLTQTSMSNMVTMTNSLSGLTLTRNVGTGGFLTTCGPLMAQQCGSQSYTTGVC 138			
QY	140 SVNSNRFSKTVAALORCOTYMDIVLVLDGSKSIYPVWEVQHFNLINLKPFYIGPGQI 199			
DB	139 SDVSDFQLRYSFAFAPQTCSFDIVVVCDENSIYPPWDARKNFLEKFVGQLDIGTKT 198			
QY	200 QGVVQYGEDVVHFFHLNDYRSVKDVAAASHIEPQRGGTETRTAFGEFARSFAQP--KG 257			
DB	199 QMGLIQAYANPRVFNLTNFKSKDEMIKATSQTFQYGGDLTNTEKAIOYARDTAYSTAAG 258			
QY	258 GRKGAKKVMVITDGESHDSDPLEKVIQOOSRDNNVTRYAVAVLYGYNRGINPETFLNEI 317			
DB	259 GRPGATKMVVVVTDGESHDGSKLKAVIDOCNKMLRFGIAVLGYARNALDKNLKIKEI 318			
QY	318 KYIASDPDDKHFNVTDEAALKDI VDALGDRIFFSLEGTKNNETSFGLMSGQTPGSSEVVE 377			
DB	319 XAIASIPYERHEFFNVSDADLLERAGTGIGEIQISIEGTIVQGDNFMQEMSVQGPSAEYSP 378			
QY	378 DG--VLLGAVGYDUNGAVLKETSAGKVIPIRESYLKEPPEL---KNHGAYLGYYTVTSV 432			
DB	379 QNNILMLGAVGAYDSMGTVQKTPGHLLI-----FSQAQFEQILQDRNHSHSYLGYSVAS- 432			
QY	433 VSRSQRYVYVAGAPFNHTGVKILFTMNHRSLTHQAMRGQQIGSYVFGSEITSVDIDGD 492			
DB	433 ISTGNSVFHVAGAPRANTGGQILVYSVNENGNTVYQSQRDQIGCSIFGSQLCAVDVXKD 492			
QY	493 GTVDVLLVGAQPYFNE-GREGRKVYVVELRNQRPVNYGTLKDSHSYNQARFGSSIASVRD 551			
DB	493 TITDVLVGAQPMYMDLKKKEGRVYLFITTKILNMMHQFLZGPNGLNENARFGSAIALSD 552			
QY	552 LNQDSYNDVVCAPLEDNHAGNIYIFPHFRGSLIKTPQRITASELA--TCLQVGCSTH 609			
DB	553 INMDGFNDVIGSPLENQSGAVIYINHEGMIRLRYSQKTLGSDRAPSSHLOQVFGSLD 612			
QY	610 GOLDLNEGDLIDLAVGALGNVILWSRVVQINASLHPFSPKINIHFROCKRSGRDATCL 669			

RX MEDLINE=98019223; PubMed=9353312;  
 RA Emsley J., King S.L., Bergelson J.M., Liddington R.C.;  
 RT "Crystal structure of the I domain from integrin alpha2beta1.";  
 RL J. Biol. Chem. 272:28512-28517(1997).  
 RN [4]  
 RP VARIANT HPA-5 (BR).  
 RX MEDLINE=94043762; PubMed=7901236;  
 RA Santoso S., Kaib R., Walka M., Kiefel V., Mueller-Eckhardt C.,  
 RT Newman P.J.;  
 RT "The human platelet allantoicins Br(a) and Br(b) are associated with a  
 RT single amino acid polymorphism on glycoprotein Ia (integrin subunit  
 RT alpha 2).";  
 RL J. Clin. Invest. 92:2427-2432(1993).  
 RN [5]  
 RP VARIANT GLU-534.  
 RX MEDLINE=20206009; PubMed=10744142;  
 RA Kroll H., Gardemann A., Fecher A., Haberbosch W., Santoso S.;  
 RT "The impact of the glycoprotein Ia collagen receptor subunit A1648G  
 RT gene polymorphism on coronary artery disease and acute myocardial  
 RT infarction.";  
 RL Thromb. Haemost. 83:392-396(2000).  
 CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,  
 CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT  
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-G-E-R IN  
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER  
 CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE  
 CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
 CC EXTRACELLULAR MATRIX.  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2  
 CC associates with beta-1. Interacts with HPSS.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -!- POLYMORPHISM: Position 534 is associated with platelet-specific  
 CC allantoicins HPA-5 (BR). HPA-5A/BR(A) has lys-534 and HPA-5B/BR(B)  
 CC has Glu-534. HPA-5B is involved in neonatal alloimmune  
 CC thrombocytopenia (NAIT or NATP). The K534E polymorphism may play a  
 CC role in coronary artery disease (CAD).  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49b entry;  
 CC WWW=http://www.ncbi.nlm.nih.gov/prow/cd/cd49b.htm".  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/  
 CC or send an email to license@ebi.ac.uk).  
 CC  
 CC -----  
 CC EMBL; X17033; CAA34894.1; -;  
 CC EMBL; AF512556; AAM34795.1; -;  
 CC PIR; A33998; A33998.  
 CC PDB; 1A0X; 25-NOV-98.  
 CC PDB; 1DZ1; 02-AUG-01.  
 CC Genew; HGNC:6137; ITGA2.  
 CC MIM; 192974; -;  
 CC GO; GO:0008305; C:integrin complex; TAS.  
 CC GO; GO:0005886; C:plasma membrane; TAS.  
 CC GO; GO:0004895; P:cell adhesion receptor activity; TAS.  
 CC GO; GO:0005518; P:collagen binding; TAS.  
 CC GO; GO:0007596; P:blood coagulation; TAS.  
 CC GO; GO:0007160; P:cell-matrix adhesion; TAS.  
 CC GO; GO:0007397; P:histogenesis and organogenesis; TAS.  
 CC InterPro; IPR000413; Integrin\_alpha.  
 CC InterPro; IPR02035; VWFA.  
 CC Pfam; PF01839; FG-GAP; 3.  
 CC Pfam; PF00357; Integrin\_A; 1.  
 CC Pfam; PF00092; vwa; 1.  
 CC SMART; SM00191; int\_alpha; 5.  
 CC SMART; SM00327; VWA; 1.

DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS00242; VWFA; 1.  
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;  
 KW 3D-structure.  
 FT SIGNAL 1 29  
 FT CHAIN 30 1181 INTEGRIN ALPHA-2.  
 FT DOMAIN 30 1132 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1133 1154 POTENTIAL.  
 FT DOMAIN 1155 1181 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 1155 1161 INTERACTION WITH HPSS.  
 FT REPEAT 45 103 FG-GAP 1.  
 FT REPEAT ? ? FG-GAP 2.  
 FT DOMAIN 188 378 VWFA.  
 FT REPEAT 378 433 FG-GAP 3.  
 FT REPEAT 434 486 FG-GAP 4.  
 FT REPEAT 488 549 FG-GAP 5.  
 FT REPEAT 551 610 FG-GAP 6.  
 FT REPEAT 615 667 FG-GAP 7.  
 FT CA\_BIND 439 507 POTENTIAL.  
 FT CA\_BIND 563 571 POTENTIAL.  
 FT CA\_BIND 527 635 POTENTIAL.  
 FT SITE 1157 1161 GEFER MOTIF.  
 FT DISULFID 83 92 BY SIMILARITY.  
 FT DISULFID 680 737 BY SIMILARITY.  
 FT DISULFID 789 795 BY SIMILARITY.  
 FT DISULFID 865 876 BY SIMILARITY.  
 FT DISULFID 1019 1050 BY SIMILARITY.  
 FT DISULFID 1055 1060 BY SIMILARITY.  
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 343 343 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 699 699 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1057 1057 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1081 1081 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 534 534 K -> E (IN ALLOANTIGEN HPA-5B;  
 FT dbSNP:1801106).  
 FT /FTID=VAR\_003977.  
 FT TURN 170 171  
 FT STRAND 173 180  
 FT TURN 183 184  
 FT HELIX 188 199  
 FT TURN 200 201  
 FT STRAND 204 204  
 FT TURN 206 207  
 FT STRAND 209 216  
 FT STRAND 220 224  
 FT TURN 226 228  
 FT HELIX 232 240  
 FT TURN 241 241  
 FT HELIX 252 262  
 FT TURN 263 284  
 FT HELIX 266 288  
 FT TURN 269 269  
 FT STRAND 275 282  
 FT HELIX 289 291  
 FT HELIX 292 301  
 FT TURN 302 303  
 FT STRAND 304 311  
 FT HELIX 313 317  
 FT TURN 318 319  
 FT HELIX 323 330  
 FT TURN 331 332  
 FT HELIX 337 340  
 FT STRAND 341 344  
 FT HELIX 347 353  
 FT HELIX 354 362  
 FT TURN 363 363  
 SQ SEQUENCE 1181 AA; 129295 MW; 7E1B7ED968A94070 CRC64;

Query Match	29.98;	Score 1863;	DB 1;	Length 1181;	
Best Local Similarity	35.66;	Pred. No. 8.1e-113;			
Matches 432;	Conservative 241;	Mismatches 433;	Indels 106;	Gaps 30;	
Qy	4	PRGLVAVWALS	WLPFGTID---TFNMWTKRPVIGPSRTAFPGYTVQQHDISENKWL	VVGA 60	
Db	10	PLPLLVLALS	-QGILNCCLAYNVGUPPEAKIFSGPSSEQGYAVQQFINFGKNWLL	VGS 67	
Qy	61	PLETNGYQKTDVYKCPV	--IHGNCYTKLNLGRVTVLSNVSRKKNMRLGLSLATNP	KDMS 117	
Db	68	PWSGFPENRGMVDVYKCPVD	LSATACEKLNLOTSIPNVTEKNTNMSLGLILTRNMG	PGG 127	
Qy	118	FLACPLWSHECGSYTTG	CMCRSVNGNFRPSKTVAPALQRCQTVMDIVIVL	DGSNIYP 177	
Db	128	PLTCGLPWAQCGNOYTTG	VCSDISDPQLSASFSPATQPCPSLIDVWVVC	DESNIYP 187	
Qy	178	WVEVQHPLNLKXFYIG	PGIOIGVYQVQGEDVHEFHLDNDRSVKDVWEAASHI	BQGG 237	
Db	188	WDVKNFLKXVQGLD	IGPTKTVGLQIYANNPRVFNLTNYTKTKKEMVAT	SQTSYGG 247	
Qy	238	TETRTAFGIBEPARSAFQ	--KGRKGAKKVMIVITDGHSDPSDLEKVIQO	SERDNVTRY 295	
Db	248	DLTWTFCAIQVARKVAYS	SAASGGRRSATKVMVVVITDGHSDGMLKAVIDC	NHDNLRP 307	
Qy	296	AVAVLGYNNRGINP	ETFLNEIKVIASDPDDKHPNVTDRALXDIVDAL	GDRFLEGT 355	
Db	308	GIADVGLNRLALOT	KMLIKKAIASIPTRYFPNVSDEALKEKAGTGL	EQIFSIHGT 367	
Qy	356	NKNETSPGLEMSQSPSSHV	--VEDGVLLGAVGAYDMNGVALKETSAGKVI	PLRESYLKE 413	
Db	368	VOGGDNFOMEMSQVGS	FADYSSQNDILMLGAVGAFGMSGTVQKTS	HGHLI----- 418	
Qy	414	FPEEL-----KXHG	VALLGYTVTVSVSSRQGRVYVAGAPFNHTK	VILTMENNRSLT 466	
Db	419	FPKQAPQIQLDRNHS	SYLGSV-AAISTGSTHFPVAGAPRANTYQG	VLVSVNENGIT 477	
Qy	467	IQHWRQOQIGSYFGS	ISITSDIDCGDVTVDLLVGAPMYFNE-GR	EKGVYVYBLQRNF 525	
Db	478	VIQHRGDQIGSYFGS	VLCSDVDVDKDTITVDLLVGAPMYMDLKE	EGRVFLTIKKGIL 537	
Qy	526	VYNGTLXDSHYQNA	RGSSTASVRDLNODSYNDVVGAPLEDHAGAI	YIFHCPGSGIL 585	
Db	538	GQHFLESGEGIE	NTFRGSAFAALSDINMOGNDVIVGSPLENQ	NSGAVIYHNGQGITR 597	
Qy	586	KTPQRRTITASBLA	-TGLQYFGCSIHGGOLDNEBDGLIDLAVCAL	GNNAVILMRFPVQINA 643	
Db	598	TKYSOKI:GSDGAF	SRHLQYFGRSLDGYDLNGDSITDVSIGAF	QVVOVLMSQSADVAI 657	
Qy	644	SLHFEPSKINIFHR	DCKSGEDATCAAFLOCTPIELAPHOTTIVG	RYNATMD----E 699	
Db	658	EASTPEKTIIVN	KNAQ-----IILKCFSAKP-RPTKNNQ	VAIVYMITLDADGFS 708	
Qy	700	RRYTPRAHLDRGG	RFTNRAYLLSSGQBLCEINRPHVLDITADY	VKPVTFVSVEYSL	EDPDH 759
Db	709	SRVTSRGLFKEN	NERCLOKQNVNQAQSCPEHI-IYIQEPS	DDVMSDLRVDISLEN	PGT 767
Qy	760	GPMLDDGMPTLR	VSVFPGWNCNDEHCVPLVDLADSRDL	PTAMEYCOVLRKPA	QDCSA 819
Db	768	SPALUAYSEAKV	SIPIPHKDCGEGDGLCISDLVDLR-QI	FAAQE-----QP----- 813	
Qy	820	YTLSPDFTVFTI	ESTRVAVEATLENRGENAYSTVLN	ISOSANLQFASLI	QKEDSDGSI 879
Db	814	-----FIVS	MNKRLETFSVTLKKNRESAYNTGIVVD	FSENLFASFS	FLPVD--GT 861
Qy	880	ECVNEERELQOV	-CNVSYPPFRAKAVAFSLDSEFSKSI	FIHLHLEILA	AGSDNSNRDS 938
Db	862	EVTQVVAASQ	SKVACDVGYPAKREQQVTFINDFNLQ	NLQNAQSL	SFOALSRSEEN- 920
Qy	939	TKEDNVAPLR	PHLYEADVLFTRSSLSHYV-----KLN	SSLEYDVGIGP	PFPCIFRIQ- 993
Db	921	-KADNLVNLK	PLLYADAEIHILTRSTNINFYBIS	DGNVPSIVHSP	EDYGPXF--IFSLKY 977



QY 120 ACSPLWSECGSYVYTCMSRVNSNFRFSKTV-A-PALQRC-QTYMDIVIVLDGNSYTP 177  
 DB 110 ADCPLSTCDQNTYLSGLCYLFRQNLQGMQGRPGQECUIGNVDVFLFDGMSLQP 169  
 QY 178 --WVEVQHFLNLIKFKYIGPQIQGVVQYGEDVYHFLNDYSVDVVEAASHIEQR 235  
 DB 170 DEFQKILDFMKDWMKK--LSNTSYQFAAVQFSTSYKTEFDFSDYVVKRKDPDALLKXVKHM 227  
 QY 236 GGTETRTAFGIEFARSEAFQK--GGRKGNKVMIVITDGEHSDSPDLKVKVQQSERDNTV 293  
 DB 228 -LLLTNTGAINVATEVFRBELGARPDATKVLIIITDGEAVDSGNIDAA-----KDLI 280  
 QY 294 RYAVAVLYNNRRGINPETFLNEIKYIASDPDKHFFNVNTE-AALKDIVDALGRIEFL 352  
 DB 281 RYIIGI-GSHFTKSSQET---LHKFASKPASE-FVKILDTFEKXLDLFTLOKXIVY 334  
 QY 353 EGTNENE-TSFGLEMSQTCFSSHVEDGVLLGAVGAYDNGAVLKETSAGKVIPLR---- 407  
 DB 335 EGTSKQDLSFNMELSSGSIADLSRGAHVAVGAKOM-----AGGLDLKADLQ 385  
 QY 408 -ESYLKEPPEELKNEHGAVLGYTVTSVSSRGREVYVAGAPRNHTGKVLFTM-----HN 461  
 DB 386 DTFPIGNEPLTEVRAGLYGYTVWLPSPQKTSILASGNAPRYOHMGRVLLFQEPGGGHV 445  
 QY 462 NRSLTIHOANRQOICGVSFSGSITSDVDIGQVTDVLLVAPMYNEGRERKVVYIELR 521  
 DB 446 SOVQTIH---GTOIGSVFGELCGVDVDQGETELELIGAPLFYGEOR-GGRVFIYQRR 500  
 QY 522 QNRVYVNGTLKDSHSONARFCSSIASVRDLNQDSVNDVVCAPLEDNHAGAIYIEHGR 581  
 DB 501 QLGFEVSELOQDPGYPIGRFEAITALTIDNGDLVDVAVGAPLEEQ--GAVYIFNHRH 558  
 QY 582 GSILTKPQRIATASLATLQVFGCSIHQGLDNLNEDGLIDLAVAGNALVILWSPPVQI 641  
 DB 559 GGLSPQPSQRIETQVLSGIQWFGERSIHGKDLGGLADVAVGAEQSMVILSSPPVDM 618  
 QY 642 NASLFEPSKINIPHRDCK-----RSGRDATCLAAFLCFTPIELAPHEQTTTV-GTRY 693  
 DB 619 VTLMSPSPABIIVHEVECSYSTSNMKKEGVNIT-----ICFOIKSLYQFQGRVANITY 673  
 QY 694 NATMDERRYTPRAHLDEGDRFTNRAVLLSSQBELCERINFIV-LDTADYKVPVFSVEY 752  
 DB 674 TLQLDGHR--TRRGLFPGGRHELNRNIAVTGMS--CTDPSFHPFVQVQDLISPINVSLNF 731  
 QY 753 SLEDDPHGP-----MLDDGHPPTILRV-----VFFWNGCNDEHCVPLVDLARSPLT 801  
 DB 732 SLWBEETPRDQRAQCKDIPPIILRPSLSHSETWEIPFKNCGDKCKEANL----- 781  
 QY 802 AMEYQVLRKPAQDCSAVTLTDFDTVFTIESTQRQVAEATLENKGENAYSTVLNISQS 861  
 DB 782 -----RVSPSPARSALRLTAFS-----LSVELSLNLEEDAYVWQDLHFFP 824  
 QY 862 ANLQFASL-IQKEDSGSIEC--VNEERLQKV--CNVSYFFFAKAKVAPRL----- 910  
 DB 825 PGLSFRKVMKPHSQIPVSCBELPEERLLSRLSCNVSSPIFKAGHSVALQMMFNTLV 884  
 QY 911 DSEPSKSLPHLHLELAAGSDNSNEEDSKENONAPLAFHLKYADVL----- 958  
 DB 885 NSWGDVSEIHL-----ANVTCNEDSOLLENSATTIPIPIPNILIQOEDSTLVVS 938  
 QY 959 FTRSSLSH-----YEVKLNS-----LERYDGIQPPFCIFRIQNLGFLPIHGM 1005  
 DB 939 FTPKPKIHQVGMVQVRIQPSIHDNHPITLEAVGVQVQ-----PSEG--- 982  
 QY 1006 KITIPATSGNRLKLRDLFTDEANTSCNINWSTETVETPE-EDLRRAPQNLHNSD 1064  
 DB 983 -----PITH-----MSVQME-PPVPCHYEDLERLUPDAABCPIC 1015  
 QY 1065 VVSINCNRILVQBNQINPHILGNLW-----RSLKALKYKSMKVMNAALQROPHSPFI 1118  
 DB 1016 GALFRCPV--VFRQILVQVICTLVLGVEIEASSMFL-CSLSISFNS--KHFF---L 1067  
 QY 1119 FREEDPSQIEPEISKQEDWQVPIWIVGSTTGGLLLLALLVLRKLGFF-RSARRRE 1177

DB 1068 YGSNASLAQVMKVVDVYVYKQM-LYLVLSGIGGULLLLLFIVLYKVGFFRNLKERME 1126  
 QY 1178 PG 1179  
 DB 1127 AG 1128

RESULT 9  
 ID ITAD\_HUMAN STANDARD; PRT; 1162 AA.  
 AC Q13349; Q15575; Q15576;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).  
 GN ITGAD.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RX MEDLINE=96111956; PubMed=8777714;  
 RA Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,  
 RA Staunton D.E., Gallatin W.M.;  
 RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-  
 RT 3";  
 RL J. Biol. Chem. 275:9959-9969(2000).  
 RN [2]  
 RP SEQUENCE OF 1-235 FROM N.A.  
 RX MEDLINE=20187620; PubMed=10722744;  
 RA Noti J.D., Johnson A.K., Dillon J.D.;  
 RT "Structural and functional characterization of the leukocyte integrin  
 RT gene CD11d. Essential role of Spi and Sp3";  
 RL J. Biol. Chem. 275:9959-9969(2000).  
 RN [3]  
 RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.  
 RX MEDLINE=96257236; PubMed=8666289;  
 RA Wong D.A., Davis E.M., LeBeau M., Springer T.A.;  
 RT "Cloning and chromosomal localization of a novel gene-encoding a human  
 RT beta 2-integrin alpha subunit";  
 RL Gene 171:291-294(1996).  
 RN [4]  
 RP INTERACTION WITH VCAM1.  
 RX MEDLINE=99059842; PubMed=9841932;  
 RA Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,  
 RA Hoffman P.A., Staunton D.E., Bochner B.S.;  
 RT "alpha beta 2 integrin is expressed on human eosinophils and functions  
 RT as an alternative ligand for vascular cell adhesion molecule 1  
 RT (VCAM-1)";  
 RL J. Exp. Med. 188:2187-2191(1998).  
 RN [5]  
 RP INTERACTION WITH VCAM1.  
 RX MEDLINE=99370002; PubMed=10438935;  
 RA Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,  
 RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;  
 RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a  
 RT binding interface between I domain and VCAM-1";  
 RL J. Immunol. 163:1984-1990(1999).  
 CC -!- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND  
 CC CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-  
 CC BORNE PATHOGENS, PARTICULATE MATTER, AND SENSIBLE ERYTHROCYTES  
 CC FROM THE BLOOD.  
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D  
 CC ASSOCIATES WITH BETA-2.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL  
 CC LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON  
 CC TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN  
 CC ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.





QY 883 --NEERLQKQVNCVYPPFRKAKVAFRLDSEFS--KSIFLHLELELAAGSDSNERDST 939  
 Db 849 VPTDEGLASSRCSVNHPIFHEGSGNCTFVTDVSVKATLGRMLMRASSENKASS 908  
 QY 940 KEDNVAPLFLHLYKADVLFTRSSLSHY-----EVKNSLSRYDGDIGPPFSCIPRI 992  
 Db 909 KA--TFQELPVPKYAVYTMISRQESSTKYFNATSDKCKMKEAERY-----RV 955  
 QY 993 QNLGLPPIHGMWKITPIATRSKNRLLKLRDPLDDEANTS--CNWNGNSTEYRPTPVEE 1050  
 Db 956 NNLQSDRL--AIGNFVFWLLNG-----VAVDVVMEAPSOQLPC-----VSEKPPQHS 1004  
 QY 1051 D----LRRAPQLNHSNDVVSNCNIRLVN---QRIHFHLLGNL---WFLSKALKYK 1099  
 Db 1005 DPLTOISRSPMLDCSTADCLQFCD---VPSFSVQBELDFTLKMLSPGWRV--ETLOKK 1059  
 QY 1100 SKIMV--NAALQKSPFIPSEDPSPQIEFISKQEDQVPIWIVGSTLGGLLALL 1158  
 Db 1060 VLVSVVAETFTQTSVTSQLPQGEAFMRAQEMVLEEDVINA--IPLIMSSVGAULLLAL 1118  
 QY 1159 LVIALRKLGFPSARRRREPGLDTPK 1185  
 Db 1119 ITATLYKLGF---KSHYKEMLEDKPE 1142  
 RESULT 10  
 ITAX HUMAN  
 ID ITAX HUMAN STANDARD; PRT; 1163 AA.  
 AC P20702;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Integrin alpha-X precursor (leukocyte adhesion glycoprotein p150,95  
 DE alpha chain) (leukocyte adhesion receptor p150,95) (CD11c) (leu M5).  
 GN ITGAX OR CD11C.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89166645; PubMed=3327687;  
 RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;  
 RT "cDNA cloning and complete primary structure of the alpha subunit of  
 RT a leukocyte adhesion glycoprotein, p150,95.";  
 RL EMBO J. 6:4023-4028(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90153906; PubMed=2303426;  
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;  
 RT "Genomic structure of an integrin alpha subunit, the leukocyte  
 RT p150,95 molecule.";  
 RL J. Biol. Chem. 265:2782-2788(1990).  
 RN [3]  
 RP ERRATUM.  
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;  
 RL J. Biol. Chem. 265:12750-12751(1990).  
 RN [4]  
 RP SEQUENCE OF 20-43.  
 RX MEDLINE=87167596; PubMed=3549901;  
 RA Miller L.J., Wiebe M., Springer T.A.;  
 RT "Purification and alpha subunit N-terminal sequences of human Mac-1  
 RT and p150,95 leukocyte adhesion proteins.";  
 RL J. Immunol. 138:2381-2383(1987).  
 CC -!- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT  
 CC RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL  
 CC INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY  
 CC IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.  
 CC -!- SUBUNIT: HETEROIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X  
 CC ASSOCIATES WITH BETA-2.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND

CC CC GRANULOCYTES.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE. family.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; M81695; AA59180.1; -;  
 CC EMBL; Y00093; CA68283.1; -;  
 CC EMBL; M29165; -; NOT ANNOTATED CDS.  
 CC EMBL; M29487; AA51620.1; ALT SEQ.  
 CC EMBL; M29482; AA51620.1; JOINED.  
 CC EMBL; M29483; AA51620.1; JOINED.  
 CC EMBL; M29484; AA51620.1; JOINED.  
 CC EMBL; M29485; AA51620.1; JOINED.  
 CC EMBL; M29486; AA51620.1; JOINED.  
 CC PIR; A36584; RWHDIC.  
 CC PDB; IN3Y; 1S-FEB-03.  
 CC Genew; HGNC:6152; ITGAX.  
 CC MIM; 151510; -;  
 CC GO; GO:0008305; C:integrin complex; TAS.  
 CC GO; GO:0004935; F:cell adhesion receptor activity; TAS.  
 CC GO; GO:0004872; F:receptor activity; TAS.  
 CC GO; GO:0007155; F:cell adhesion; TAS.  
 CC GO; GO:0007337; P:histogenesis and organogenesis; TAS.  
 CC InterPro; IPR000413; Integrin\_alpha.  
 CC InterPro; IPR002035; VWF\_A.  
 CC Pfam; PF01839; FG-GAP; 3\_  
 CC Pfam; PF00357; integrin\_A; 1.  
 CC Pfam; PF00092; vwa; 1.  
 CC PRINTS; PRO1185; INTEGRINA.  
 CC PRINTS; PRO0453; VWFADOMAIN.  
 CC SMART; SM00191; Int\_alpha; 5.  
 CC SMART; SM00327; VWA; 1.  
 CC PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 CC PROSITE; PS50234; VWFA; 1.  
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 CC Signal; Magnesium; Calcium; Repeat; 3D-structure.  
 CC SIGNAL 1 19  
 FT CHAIN 20 1163 INTEGRIN ALPHA-X.  
 FT DOMAIN 20 1107 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1108 1128 POTENTIAL.  
 FT DOMAIN 1129 1163 CYTOPLASMIC (POTENTIAL).  
 FT REPEAT 34 87 FG-GAP 1.  
 FT REPEAT ? ? FG-GAP 2.  
 FT DOMAIN 165 351 VWFA.  
 FT REPEAT ? ? FG-GAP 3.  
 FT REPEAT 402 453 FG-GAP 4.  
 FT REPEAT 455 517 FG-GAP 5.  
 FT REPEAT 518 576 FG-GAP 6.  
 FT REPEAT 581 633 FG-GAP 7.  
 FT CA\_BIND 466 474 POTENTIAL.  
 FT CA\_BIND 530 538 POTENTIAL.  
 FT CA\_BIND 593 601 POTENTIAL.  
 FT SITE 1131 1135 GFFKR MOTIF.  
 FT DISULFID 69 76 BY SIMILARITY.  
 FT DISULFID 108 126 BY SIMILARITY.  
 FT DISULFID 655 712 BY SIMILARITY.  
 FT DISULFID 771 777 BY SIMILARITY.  
 FT DISULFID 848 863 BY SIMILARITY.  
 FT DISULFID 998 1022 BY SIMILARITY.  
 FT DISULFID 1027 1032 BY SIMILARITY.  
 FT CARBOHYD 61 61 N-LINKED (GLCNAC... ) (POTENTIAL).

FT CARBOHYD 89 N-LINKED (GLNAC. . .) (POTENTIAL).  
FT CARBOHYD 392 N-LINKED (GLNAC. . .) (POTENTIAL).  
FT CARBOHYD 697 N-LINKED (GLNAC. . .) (POTENTIAL).  
FT CARBOHYD 735 N-LINKED (GLNAC. . .) (POTENTIAL).  
FT CARBOHYD 899 N-LINKED (GLNAC. . .) (POTENTIAL).  
FT CARBOHYD 939 N-LINKED (GLNAC. . .) (POTENTIAL).  
FT CARBOHYD 1050 N-LINKED (GLNAC. . .) (POTENTIAL).  
FT CONFLICT 490 G -> A (IN REF. 2).  
FT CONFLICT 756 L -> D (IN REF. 2).  
FT SEQUENCE 1163 AA; 127885 MW; 6C4E19CC3P62A473 CRC64;  
  
Query Match 18.2%; Score 1132.5; DB 1; Length 1163;  
Best Local Similarity 28.4%; Pred. No. 2.1e-65;  
Matches 354; Conservative 219; Mismatches 471; Indels 201; Gaps 48;  
  
QY 7 LVVAVALSIPGFTDFNMDTRKPRVPIGSRAP-----FGVTVQOHDISGNKVLVGA 60  
DB 10 LFTALATSL-----GFNLDTEB-----LFAFVDSAGFGDSVVQY---ANSVVVGA 53  
QY 61 PLETNQYKTDYKCPVHNGCTKLNLRVTSNYSERKDNRLGLSLATNPKDNSFLA 120  
DB 54 POKITAAQTGLGYQCYGTGACEPTGL-----QVPEAVNMSLGLSLASTTSPQLLA 107  
QY 121 CSPLSHSCGSSYYITGMSRVNSNFRFSKTVAPALQRC-QTYMDIVIVLDSGSIYP--- 177  
DB 108 CGPTVHEGCRNMVLTGLCLLGT-QLTORLPVSRQECPRQEQDIVFLIDGSGSISRN 166  
QY 178 WVEVQFLINILKFKFYIGPQQVGVQYCEDVHEFHLDNYSRVKDVVVEASHIEQ-RG 236  
DB 167 FATMNFVRAVISQFQ--RPSQFSLMQFSNKFQCHFTFEERFRTSNPLSLLASVHQLQ 224  
QY 237 GTETRAPGIBPARSAFOGKRGKAKKVMIVITGESH-DSFDLEKVIQOGERNVTRY 295  
DB 225 FVTATAIQVHRLPHASYGARRDATKILVITGKGGSLDYKDVIPMDAAGITRY 284  
QY 296 AVAV-LGYNNRGINPETELNEIKYIASDPDDKHPNVDEAALQIDVADGRIFSLG 354  
DB 285 AICVGLAFQNRNS-----WKELNDIASKPSQEHFKVEDPDAIKDIONQKEKIFATG 338  
QY 355 T-NKNETSPGLEMSQTFGSHVVEDGVILGAVGAYDNGAVIKETSAKVIP--LRESYL 411  
DB 339 TETTSSSFELEMAQGFSAVTPDPGVILGAVGSPFTWSGAF-----LYPPNMSPTFI 391  
QY 412 KEPPBELKHGAVLGYTVTSVSSRQGRVYVAGAFENHTGVILFTWHNNESLTHQAM 471  
DB 392 NMSQENVDRSDSLGYS--TELALWKGVSILVQAPRYQHTGAVIPT-QVSRQNRKAEV 449  
QY 472 RGQIGSYFGSEITSYDIDGDGVTDLVLLGAPMYFNEGERGKVYYELRQ--NRFYNG 529  
DB 450 TGQIGSYFGASLCSVDVDTGSDTLVILGAPHYBQTR--GGQVSVCLPRGWRWACDA 508  
QY 530 TLKDSHYQYARFGSSIASVRDLNODSVYDVVVGAPLEDNHAGAVIHPGFG-SILKTP 588  
DB 509 VLGSGQHPWGRFGAALTVLGDVNGDKLTDVWIGAPGEBENRGAVLPHGVLPSPISPSH 569  
QY 589 KQRTASELATGLQYFGCSHIGOLDLNEDEGLDILAVGALGNVILMSRPPVQVQINASLHPE 648  
DB 569 SQRIAGSLSRLQYFGQALSGQODLTQGLDVLAVGARGQVLLLRTPVLVWGVSMQFI 628  
QY 649 PSKINIFHRDKES--GRDATCLAAFLCF-----TFIPLAFPTTTCVIRYNATDRR 701  
DB 629 PAEIPSAFECRQVSVSEOTVQSNICLYIDKRSNVLGSRDLQSS-----VTLDLALDPCR 685  
QY 702 YTPRAHLDGGDRFTNRVALLSGQBLCEIRINPHVL---DTADYKVPVTFVSVEYSLEDP- 757  
DB 686 LSPRAFQETKNSLSRVRVIGL-KAHCE--NFNLLPSCVDSVTPIRLNFTLVGKP 742  
QY 758 -----DHGPMLDGQWTTTTRVSVFPWNGCNEDHCVPDVLDRSDLPATMEYCORVLK 812  
DB 743 LLAFLRLPMLAALAQRYFTASLPFKXCGADHIC----- 777  
QY 813 PAODCSAYLSFDTTTFTIESTQRVAVETALNENGENAYSTVLNISQANLQFASLI-- 870

DB 778 --QDNILGFSFPGKLSLLVGSNLENAEVMVMDGSDSYGTTTTFSHPGAGLSYRYVABG 835  
QY 871 QKESDGSIGECVNEBRLOKO-----VCNVSYPFRAXAKVAFRLDSEFS-KSIFLHLE 924  
DB 836 QKQQLSLHLTCDTAPVSGSQSTWSTSCRINHELIFRGAQITFLATPVPSPRAVLGDRLL 895  
QY 925 IELAAGSDSNRSDTKEDNVAFLPHLYKYEADVLFTTRSSLSHSHVEVKLNS----- 975  
DB 896 LTANVSSENNTPTSK--TTFQELFPVKY---AVTVVSSHEQTKYLNFSSEKEKSHV 950  
QY 976 -LERYDGIGPPFCIFRIQNLGL--FPIHGMMKITTPIATRSNRLKLKLDFTLDEANT 1032  
DB 951 AMHYR-----QVNNIGQDLFPV---SINFWPVE-----LNQEA-- 981  
QY 1033 SCNTWGN-----STEYRPTPVE--DLRRAPOLNHSNDVVSINCNIRLVP 1076  
DB 982 ---VMDVEVSHQPQNSLRCSEKIAFPASDFAHIQKPNVLDCSIAGCLAFRC--VP 1035  
QY 1077 N-----QEIFHLNL---WLSLKALKYKSM---KIMVNAALQRFHSPIFREDDSR 1126  
DB 1036 SFSVQBEHLDFTLGNLSFGWVRQILQKKYSVSVABITFTDTSVYSQLPQGFQAFVRAQTT 1095  
QY 1127 QIB-PEISKQEDMQVPIWIVGSTLGGLLALLLALLVLRKLGFFR 1170  
DB 1096 VLEKIKVHN-----PTPLVSGSSIGLLALLLITAVLYKVGFFK 1134  
  
RESULT 11  
ITAM\_MOUSE STANDARD; PRT; 1153 AA.  
AC POS555;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-FEB-1991 (Rel. 17, last sequence update)  
DT 21-FEB-2003 (Rel. 41, last annotation update)  
DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1).  
OS ITGAM.  
GN Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89312584; PubMed=3044779;  
RA Pyela R.;  
RT "Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor."  
RL EMBO J. 7:1371-1378(1988).  
RN [2]  
RP SEQUENCE OF 11-45 FROM N.A.  
RC STRAIN-BALB/c; TISSUE=Spleen;  
RX MEDLINE=86287312; PubMed=2942940;  
RA Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E., Larson R.S., Roberts T.M., Springer T.A.;  
RT "A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 3 and cellular adhesion molecule Mac-1.";  
RN Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).  
RN [3]  
RP SEQUENCE OF 17-28.  
RX MEDLINE=8518276; PubMed=388182;  
RA Springer T.A., Teplow D.B., Dreyer W.J.;  
RT "Sequence homology of the IFA-1 and Mac-1 leukocyte adhesion glycoproteins and unexpected relation to leukocyte interferon."  
RL Nature 314:540-542(1985).  
CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. ALPHA-M/BETA-2 PLAY A CRITICAL ROLE IN



Db 783 TMSAGLDT---LVVGGPQDFNMSVTLRNDGBDSYGTQVTVVYPSGLSYRKDSASQNP 839  
QY 869 -----LQKEDSDSIECWNERBLQKQVCNVSPPFRAKAKVAFRLDSEF-SKSIPLHL 923  
Db 840 KKPWFVKPAPSSSSSE---GHGALKSTTWNINHPFPANSEVTFTVDVSHASFGNKL 896  
QY 924 EIELAAGSDNERDSTKEDNVAPLRFHLKYEADVLFTRSSSLSHY-----EVLKSS 975  
Db 897 LKAIIVASENM---SRHTKTKFQLELPVKYAIWTVTSDSSIRVYNFTASEMTSKVICH 954  
QY 976 LERYDSIGP---PFCIP-----RIONLGLPPHGGMMKTIPIATSGNRLKLRLD 1028  
Db 955 QYQFNELGQSLVSVVFPVQINNVTWD-HQVI-----PQ 993  
QY 1029 EAMTSCNMGNSTEYRTP---VEEDLRAPOLHNSDVWSINCIRLVPNOEI-NFHL 1084  
Db 994 NLSSACH-----TEQSPHNSFRDQLERTFVLNCSVAVCKRIQCDLPSPNTOEINVT 1048  
QY 1085 LGMH---WLRLK---ALKYKMKIMVNAALQRFSPFTFREEDPSROIETFSIQEDW 1138  
Db 1049 KGNLSPDWYIKTSHGHLVSSTEILFN-----DSAFALLQGSYVRSKRTETKVPY 1101  
QY 1139 QV---PTWIIIVGSLGLLMLLALVLRKLGFPR 1170  
Db 1102 EVHNPVLVIGSSIGGLVLLALITAGLYKLGFPR 1135

RESULT 12

ITAM\_HUMAN STANDARD; PRT; 1152 AA.  
AC P11215;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MOI) (Neutrophil adherence receptor).  
GN ITGAM OR CR3A OR CD11B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RX MEDLINE=88315033; PubMed=2457584;  
RA Corbi A.A., Kishimoto T.K., Miller L.J., Springer T.A.;  
RT "The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";  
RL J. Biol. Chem. 263:12403-12411(1988).  
RN [2]  
RX MEDLINE=88190151; PubMed=2833753;  
RA Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;  
RT "Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein MOI: chromosomal localization and homology to the alpha subunits of integrins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).  
RN [3]  
RX MEDLINE=88257215; PubMed=2454931;  
RA Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;  
RT "Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor MOI (complement receptor type 3).";  
RL J. Cell Biol. 106:2153-2158(1988).  
RN [4]  
RX MEDLINE=93123748; PubMed=8419480;  
RA Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;  
RT "Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";

J. Immunol. 150:480-490(1993).  
RN [5]  
RX MEDLINE=89098993; PubMed=2563162;  
RA Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;  
RT "cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).  
RN [6]  
RX MEDLINE=92073318; PubMed=1683702;  
RA Shelley C.S., Arnaout M.A.;  
RT "The promoter of the CD11b gene directs myeloid-specific and developmentally regulated expression.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).  
RN [7]  
RX MEDLINE=92144986; PubMed=1346576;  
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;  
RT "Characterization of the myeloid-specific CD11b promoter.";  
RL Blood 79:865-870(1992).  
RN [8]  
RX MEDLINE=87076671; PubMed=3539202;  
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;  
RT "N-terminal sequence of human leukocyte glycoprotein MOI: conservation across species and homology to platelet IIb/IIIa.";  
RL Biochim. Biophys. Acta 874:368-371(1986).  
RN [9]  
RX MEDLINE=95171458; PubMed=7867070;  
RA Lee J.O., Rieu P., Arnaout M.A., Liddington R.;  
RT "Crystal structure of the A domain from the alpha subunit of integrin CR3 (CD11b/CD18).";  
RL Cell 80:631-638(1995).  
RN [10]  
RX MEDLINE=96363671; PubMed=8747460;  
RA Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;  
RT "Two conformations of the integrin A-domain (I-domain): a pathway for activation?";  
RL Structure 3:1333-1340(1995).  
RN [11]  
RX MEDLINE=98362595; PubMed=9687375;  
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A., Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrichson R.L., Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E., Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;  
RT "Cation binding to the integrin CD11b I domain and activation model assessment.";  
RL Structure 6:923-935(1998).  
RN [12]  
RX MEDLINE=98226734; PubMed=9560195;  
RA Oxvig C., Springer T.A.;  
RT "Experimental support for a beta-propeller domain in integrin alpha subunits and a calcium binding site on its lower surface.";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).  
CC -1- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES. IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN.  
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M ASSOCIATES WITH BETA-2.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND



781 ----SITFSFMSLCLVVGPPREFNVTVTVVRNDCGDSYRTQVTFPPDLISYKVKSTLQW 836  
873 EDSXGS--IECVNEER-----RLQKQVCNVSYPFRKAKAVARLDSER-SKSIPLHLHL 924  
837 QRORSWRLACESASSTEVSGALKSTCSINHIFPENSVTNITPVDVSKASLGNKLL 896  
925 IELAAGSDNSRSTKEDNVAPIRHLKYEADVLFTRSSLSHYEYKLNSSLERYDGI 984  
897 LKANVTSENPNRNTKTE--FQLELPVKYAVYVMVTVSHGVSTKY--LNFTAS--ENVIS 949  
985 PFCIFELQNLGLFPIHGMWMTIPIATSGNRLKLRDLFLDEANTSCNIWGNSTEYR 1044  
950 VMQHQYQVSNLQ---RSLPISLVFLVFLVNLQTVIWDREPVTFSENLSSTC--HFKERL 1004  
1045 PTPVE--EDLRRAPOLNHSNDSVVSINCLRLVP--NOEINPHLLGNL---WLSLX---A 1095  
1005 PSHSDFELAEUKAPVNVCSIAVCORIQCDIPFFGIQIEFNATLKNLSFDFWYKTSNHL 1064  
1096 LKYSKMKIMNAALQROFHSPFFREDPSRQIE--EISKQEDQWQVIMLIVGSTIGELL 1154  
1065 LIVSTAEILFNSVFTLLPQGGAFVRSQTETKVEPEVEN-----PLPLIVGSSVGGELL 1118  
1155 LLALLVLALRLKGLFFRSARRRE-----PGLDP 1182  
1119 LLALITAALYKLGFEK--RQYKDMWSGGPPGAEP 1151

RESULT 13  
ITAL MOUSE  
ID ITAL MOUSE STANDARD; PRT; 1163 AA.  
AC P24063;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE integrin alpha-L precursor (leukocyte adhesion glycoprotein LFA-1  
DE alpha chain) (Leukocyte function associated molecule 1, alpha chain  
DE (CD11a).  
GN ITGAL OR LFA-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91268576; PubMed=2051027;  
RA Kaufmann Y., Tseng E., Springer T.A.;  
RT "Cloning of the murine lymphocyte function-associated molecule-1  
RT alpha-subunit and its expression in COS cells."  
RL J. Immunol. 147:363-374 (1991).  
RN [2]  
RP SEQUENCE OF 24-42.  
RX MEDLINE=85188276; PubMed=3887192;  
RA Springer T.A., Teplow D.B., Dreyer W.J.;  
RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion  
RT glycoproteins and unexpected relation to leukocyte interferon."  
RL Nature 314:540-542 (1985).  
CC -1- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,  
CC ICAM3 AND ICAM4. IS INVOLVED IN A VARIETY OF IMMUNE PHENOMENA  
CC INCLUDING LEUKOCYTE-ENDOTHELIAL CELL INTERACTION, CYTOTOXIC T-CELL  
CC MEDIATED KILLING, AND ANTIBODY DEPENDENT KILLING BY GRANULOCYTES  
CC AND MONOCYTES. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-L  
CC SUBUNIT GENE DEMONSTRATE IMPAIRED TUMOR REJECTION AND IMPAIRED  
CC LEUKOCYTES RECRUITMENT.  
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-L  
CC ASSOCIATES WITH BETA-2.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: LEUKOCYTES.  
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -1- SIMILARITY: Contains 1 VWFA domain.  
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).

-----  
EMBL; M60778; AAA39426.1; .  
PIR; I56126; I56126.  
DR HSSP; P20701; ILPA.  
DR MGD; MGI:96506; Itgal.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01939; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PRO1185; INTEGRINA.  
DR PRINTS; PRO0453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS0234; VWFA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Signal; Magnesium; Calcium;  
KW Repeat.  
FT SIGNAL 1 23  
FT CHAIN 24 1163  
FT DOMAIN 24 1084  
FT TRANSMEM 1085 1108  
FT DOMAIN 1109 1163  
FT REPEAT 39 88  
FT REPEAT 7 77  
FT DOMAIN 148 334  
FT REPEAT 399 454  
FT REPEAT 455 514  
FT REPEAT 516 573  
FT REPEAT 576 628  
FT CA\_BIND 466 474  
FT CA\_BIND 528 536  
FT CA\_BIND 588 596  
FT SITE 1111 1115  
FT DISULFID 70 77  
FT DISULFID 108 126  
FT DISULFID 147 199  
FT DISULFID 651 705  
FT DISULFID 767 773  
FT DISULFID 840 856  
FT DISULFID 993 1009  
FT DISULFID 1017 1048  
FT CARBOHYD 86 86  
FT CARBOHYD 185 185  
FT CARBOHYD 270 270  
FT CARBOHYD 444 444  
FT CARBOHYD 668 668  
FT CARBOHYD 696 696  
FT CARBOHYD 724 724  
FT CARBOHYD 728 728  
FT CARBOHYD 776 776  
FT CARBOHYD 857 857  
FT CARBOHYD 880 880  
FT CARBOHYD 890 890  
FT CARBOHYD 899 899  
FT CARBOHYD 927 927  
FT CARBOHYD 1056 1056  
SQ SEQUENCE 1163 AA; 128343 MW; A7A3078489B8232F CRC64;

Query Match 17.2%; Score 1073.5; DB 1; Length 1163;  
Best Local Similarity 28.1%; Pred No. 1.4e-61;  
Matches 358; Conservative 206; Mismatches 450; Indels 261; Gaps 55;

4 PRGLVAVWAL--SLWPGFTDFNMDTRKPRVPGSRTAFFGYTVQOQHDISGNKMLVWGAP 61

Db 8 PRULLLGLQFAKAM-----SNLDRPTQSFIAQGRHFGYQVLOIE-----DGVVVGAP 58  
Qy 62 LETNGYQKTDGDKYKVIHGNCTKMLNGLRVTLNSVSEKXNMRGLGLSLATNPKDMSFLAC 121  
Db 59 GEGD---NTGLYHCHRTSEFCOPVSLH-----GSNHTSKYLEMTLATDAAGSLAC 108  
Qy 122 SPLWSECCSSYVTCMCVRNSNFPKTV-----APALQRC-QTYMDIVIVLDGSL 173  
Db 109 DGLSRECDQNTYLSGLC-----YLFQSLBGPMLQNPAYQCECMKGVDVLFVDFGSG 162  
Qy 174 SI--YPMVBOVQFLINILKFKYIGPQIQWGVVQGEDVVEFHLDNDRSVK-----DVV 226  
Db 163 SLDRKDFEKILEFMKDMVK-----LSNTSYQFAVQSDCTCTETFLDY--VKQKNPQV 218  
Qy 227 BAASHIEQGGTE-----TRTAPGIEPARSEAP--OKGORKAKKVMIVITGESHSDPD 279  
Db 219 -----LGSVQPMFLANTFRAINVYVAHFKEESGARPDATKVLVIITGGEASDGN 270  
Qy 280 LEKVIQSSRDNVTRVAVAVLGYNNRRGINPTEFLNEIKYIASDPDDKHFFNVDE--AAL 338  
Db 271 I-----SAADITRIIGIKHF--VSQKQKTLH---IFASEPVEE-FVKILTTFEKL 318  
Qy 339 KQIVDALGDRIFSELTNNKE--TSFGLNSQTFSSHVVDGVLGAVGAYDNGAVLKB 397  
Db 319 KOLFDTLQRIYAIEGTNRQDLTSFNNELSSGISADLSKGEAVVGAAGAKDW----- 371  
Qy 398 TSAGKVIPIRE-----SYLKKEPEELKNEGAYLGYTVTSVVSQRQGVVAGAPRNFHTG 452  
Db 372 --AGGFEDREDLQGAFTVQOEPLTSDVGGYLGTVVAMWTSRSPRLAAGAPRQHV 429  
Qy 453 KVLFTMHN-----NRSLSITHAMRQCGQSGFSGSEITSDVIDGDGVTDLVLVAGAPYFN 507  
Db 430 QVLLFOAPEAGGRWNT-----QKIEGTQSGYFGGELCSVDLDQDGEABELLIGAPLFG 485  
Qy 508 EGBERKVVYELRQNFYNGTILKDSHYSQNRFGSSIASVRLDNLQDSYNDVVVCAPIE 567  
Db 486 EQR--GKRVTYQRQLFEMVWSLQDGPYVGRFGAAITALTIDNGDRLTDVAVGAPLE 544  
Qy 568 DNHAGIYIFHGRGSIKTPKQITASBLATGLQYFGCSINGQLDNLNEDGLIDLAVAL 627  
Db 545 EQ--GAVYIFNGKPGGLSPQSQIOGAQVFFGIRWFGSIHGVGLDGLDLADVVGAE 602  
Qy 628 GNAVILWSPVVOINASLHPSKINIHRDCKRSGRDATCLAAFL--CFTPIFLAPHPQ 685  
Db 603 GRVVLSRPVVDVVTLSFSPBIEFVHEVCSYSARBEQKGVKLKACFRKPLTPPQ 662  
Qy 686 -TTVGIRYNATWDERRYTPRAHLDGGDRFTNRAVLLSSGQFLCERINPHV-LDTADV 743  
Db 663 GRILLANLSTQLDCHRMBSRGLPPDGSHLSQNTSI--TPKXSLDFHFPICIQDLI 720  
Qy 744 KPVTFSEVYSLEDDHGMMLDDG-----WPT--TLRVSVVPMWGCNEDHCVPDLVLD 794  
Db 721 SPINVSINFTSLLEEGTPRDQGRAMQPILRPSHTVTKRIPPEKNGEDKXCEANLJLS 780  
Qy 795 --ARSDLPAMEYQVRKPAQDCSAYTLSPDTTFFIESTQRVAZEATLENENAY 852  
Db 781 SPARG-----PURLSSASL-----AVEWTLNSGEGDY 810  
Qy 853 STVLNISQSANLQF--ASLIQ-----KEDSDGSIECVAEERLQOV--CNVSYPPF 900  
Db 811 WVELDLPFRGLSFRXVEMQLPHSRMPVSCBELTEGS-----SLLTNKLKCNVSSPIF 863  
Qy 901 RAKAKVAPR-----LDSEFSKIFLH---HLEIELAAGSDSNERDSTKEDVAPRFLH 951  
Db 864 KAGQEVSLQVMNTLLNSWEDFVGLNGTWKCE-----NENSSLOQEDNSAATHPV 914  
Qy 952 KYEADVL-----FT-----RSSLSLH--YEVKLNS-----LEYDGIQGPFP 986  
Db 915 LYPVNLITKEQENSTLYISFTPKGPKTQVQVHYQVRIQPSAYDHNMPTLEALVGVPRH 974  
Qy 987 SCIFRIONLGLFPPIHGMKMTIPIATRSGRNLLKLRLDPLTDEANTSCNIGWNTSTYRPT 1046

975 S-----EDLIT-----YTKSVQTDPLVT 992

1047 PVEEDLRAPQANHSNDVSVSINCINRLVNPQNEINHLNML--RSKALKYKSM--KI 1103

993 CHSEDLKR-PSSEAEQPCLPQVPRCPVPRWEILQVGTVELSKETKASSTLSLCSL 1051

1104 MVNAALQRPSPPIFREEDPSROIPEISQEDQVPIWIVGSLGALLLALLLAL 1163

1052 SVFSNNSKHFF--LYGSKASEAQVLVKVDLHKEM-LHVTIVLSGIGLVLFLFLAL 1107

1164 RKLQFF-RSARRRE 1177

1108 YKVGFFKXNLKEME 1122

RESULT 14

ITAE MOUSE

ID ITAE MOUSE STANDARD; PRT; 1167 AA.

AC Q60677;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Integrin alpha-E precursor (integrin alpha M290).

GN ITGAE.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AKR;

RX MEDLINE=95187992; PubMed=7882170;

RA Smith T.J., Ducharme L.A., Shaw S.K., Parker C.M., Brenner M.B.,

RA Kileshaw P.J., Weis J.H.;

RT "Murine M290 integrin expression modulated by mast cell activation.";

RL Immunity 1:393-403(1994).

CC -!- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT

CC MEDIAL ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL

CC CELL MONOLAYERS. MICE EXPRESSING A NULL MUTATION OF THE ALPHA-E

CC SUBUNIT GENE EXHIBIT A MARKED REDUCTION IN THE NUMBERS OF

CC INTRAEPITHELIAL LYMPHOCYTES IN THE GUT AND IN THE DEVELOPMENT OF

CC GUT-ASSOCIATED LYMPHOID AGGREGATES, SUPPORTING A SPECIFIC ROLE FOR

CC THIS INTEGRIN IN MEDIATING RETENTION OF LYMPHOCYTES IN THE

CC INTESTINAL WALL.

CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA

CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A

CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS

CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

CC -!- SIMILARITY: Belongs to the integrin alpha chain family.

CC -!- SIMILARITY: Contains 1 VWFA domain.

CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U12236; AAC52142.1; ..

DR HSP; P11215; IABX.

DR MGD; MGI:1298377; Itgae.

DR InterPro; IPR000413; Integrin\_alpha.

DR InterPro; IPR002035; VWF\_A.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00357; Integrin\_A; 1.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int\_alpha; 3.



DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN ALPHA; 1.  
DR PROSITE; PS0234; VWFA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Signal; Repeat; Magnesium;  
KW Calcium.  
FT SIGNAL 1 19 BY SIMILARITY.  
FT CHAIN 20 1167 INTEGRIN ALPHA-E.  
FT CHAIN 20 181 INTEGRIN ALPHA-E LIGHT CHAIN.  
FT CHAIN 183 1167 INTEGRIN ALPHA-E HEAVY CHAIN.  
FT DOMAIN 20 1114 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 1115 1137 POTENTIAL.  
FT DOMAIN 1138 1167 CYTOPLASMIC (POTENTIAL).  
FT REPEAT ? ? FG-GAP 1.  
FT REPEAT ? ? FG-GAP 2.  
FT DOMAIN 149 192 X-DOMAIN (EXTRA DOMAIN).  
FT DOMAIN 193 384 VWFA.  
FT REPEAT ? ? FG-GAP 3.  
FT REPEAT 449 501 FG-GAP 4.  
FT REPEAT 503 564 FG-GAP 5.  
FT REPEAT 566 631 FG-GAP 6.  
FT REPEAT 634 686 FG-GAP 7.  
FT CA BIND 514 522 POTENTIAL.  
FT CA BIND 578 586 POTENTIAL.  
FT CA BIND 646 654 POTENTIAL.  
FT DOMAIN 185 191 GLU-RICH (ACIDIC).  
FT SITE 1140 1144 GPKR MOTIF.  
FT DISULFID 72 83 BY SIMILARITY.  
FT DISULFID 130 164 BY SIMILARITY.  
FT DISULFID 698 754 BY SIMILARITY.  
FT DISULFID 814 820 BY SIMILARITY.  
FT DISULFID 898 898 BY SIMILARITY.  
FT DISULFID 998 1023 BY SIMILARITY.  
FT DISULFID 1031 1047 BY SIMILARITY.  
FT CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 256 256 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 314 314 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 718 718 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 773 773 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 829 829 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 846 846 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 925 925 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 968 968 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1013 1013 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1086 1086 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 1167 AA; 128983 MW; B8331C115DCCCFD CRC64;

Query Match 15.3%; Score 954; DB 1; Length 1167;  
Best Local Similarity 26.2%; Pred No. 8.2e-56;  
Matches 340; Conservative 203; Mismatches 459; Indels 298; Gaps 51;  
QY 10 AWALSPLPGFTDFNMTRKPRVPGSRTPAPGYTVOQHDHISGNKWLWVGCAPLETNGYQK 69  
DB 27 AWVTAQPG-----APAVLS-----SLLHLDPSNNQCLLVARRSSN--RN 65  
QY 70 TGDVYKC-----PVHGNCTKLNGLRVTLSNVSRKDNRLGLSLATWPKNS 117  
DB 66 TAALRYCAISIPDEACQVEHCMPK---GRYQ-----GVTLVGN--HNG 107  
QY 118 FLACSPLMWSECGS-SYTTTGMCSRVNSNF----- 146  
DB 108 VLVCIQVQARKPRSLNSLNGACSLTPNLDLQAQVAFSDLEGFLDQPAHVDSDGYCRSK 167  
QY 147 -----RSKTVAPALQRCQTMDIVLDGNSNIYP--WVEVOHFLINILKXKY 193  
DB 168 GSGTGEKKSARRRTVEEDBE-EDGTEIAIVLDGSGSIGSDPQKAKNFISTMWRNFY 226

QY 194 IGPQIQVGVVOYGEDVVEHFLNDYRSVKDVVVEAASHIEHQGGTGTETRTAPGIEFARSA 253  
DB 227 EKCFCFNALVOYGAVIQTEFDLQSRDINASLAKVQSIQVQKEV-TKTASAMQHVLDMI 285  
QY 254 F--QXGGRKGAKKMIVITDGESHDP--DLKVIQOOSERDNTVTRAVAVLGYNNRGINP 310  
DB 286 FIPSGSRKALKVMVLTGDIQDPLNLTIVINSPKQGVVRFAGV---GRFRKFN 341  
QY 311 EYFLNEIKYIASDPDDKHPFNVTDBAALKDVIDALGDRIFFSLEGTKMNETSFCLEMSQTS 370  
DB 342 NTY-RELKLIASDPKEAHTPKVNTYALDGLLSKLQQRIVHMEGTVGDAQY--QLAQITG 398  
QY 371 FSHHVDEG-VLLGAVGAYDN--GAVLKETSAGKVIPLRESYLKFPPELKN--HGAYLIG 426  
DB 399 FSAQILDKQVLLGTGVAFFNMSGGALLYSTONG-----RGCFLNQAKEDSRTVQVSYLG 453  
QY 427 YVTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNNRSLITHOAMRGQIQGYGSEITS 486  
DB 454 YSL-AVLHKAHGISYVAGAPRHLKAGAVFELKXEDREDAFVRRIEGQMGSYFSGVLCF 512  
QY 487 VDIQDGVTVLLVGAPMYFNREGRGKVVYVELRQN--RFVYNGTLKDSHSHYQNAFQS 544  
DB 513 VDIQDGVTVLLVAAFPYHIRG-BEGRVYVYVQVPEQDASFSLAHTLSGHPGLTNSRFGF 571  
QY 545 STIASVRDLNQDSYNDVVVGCAPLEDNHA-----GAIYIFHGPRGSIKTPKORITASELA 598  
DB 572 ANAAVGDINQDKFTDVAIGAPLEGAGDGASGVYIYNGHSGHLYSPSQIIRASSVA 631  
QY 599 TGLQYFGCSIHQDLNEDGLIDLAVGALGNVILMSRPVQVAINASLHFEPS----- 650  
DB 632 SGLHYFGMSVSGGLDFNGDGLADITVGSRDSAVLRSRPVVDLTVSMFTPDALPMVFTG 691  
QY 651 --KINI--PHROCKRSRGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYTPRAH 707  
DB 692 KMDVNLCEVDGSSVVASPEGLREMFNFT-----VDVDTKORQLQ 733  
QY 708 LDE-----CGDRFTNRAVLLSSGQELCEINRPHVLDLTADYVVPVTSVYSLED 756  
DB 734 CEDSSGQCLRKWNGSFLCEHFWL1STEELCE-----DCFSNITIKTYVEFOT 784  
QY 757 P----DH--GPMDDGWPITLAVSVFPWNGCNEDEHCVLPDLVDARSDLPTAMEYCORVLR 811  
DB 785 SGRRDYFNPLDHYKEPSAIPOLPYEKCKNKFCAIQL----- 826  
QY 812 KPAQDCSAYTLSDFTTVFIIESTRQVAVEATLENRGENAYSTVLNISQSANLQFASLIQ 871  
DB 827 -----TTWISQQLVGVYKX--VTMNIISLTNSGDSYMTNMAVPRNLQPKK-IQ 875  
QY 872 KEDSGSIECVNEERRLOQV--QNVSPYPPRAKAKVAPRLDSEFSKSIPLHLEIELAA 929  
DB 876 REVSP-DVQCDQDPKPVASVLMNCKIGHPIPK-RSSVNVSVTWQLEESVFPNRTADITVT 933  
QY 930 GSDSNRSDTKEDNVAPLRPHLKVEA-----DVLFP--TRSSSLSHYVKLSSLSLRYDGI 982  
DB 934 ISNSNEKSLAR--TRSLQFPHAFIATVLSRPSVMYMTSQSPSDHKEFFENVHGENL--- 988  
QY 983 GPPFCIFRIQNLGFLPIHGMMKITIPIATRSGNRLKLKRDFTLDEANTSCNIWGNSTE 1042  
DB 989 ---FGAVFQLQ-----ICVPIQLQD-FQIVRVKNLTKTDQTEC----- 1023  
QY 1043 YRPTVEEDLARAQNLN-----SNSDVVSINCMRLVPMQOEINFHLLGNLWL 1090  
DB 1034 ---TOSQBPACSDPVQVQVKEHWSVVCATSNKENVTVAELSVGHTKQL-----L 1071  
QY 1091 RSLKALKY-----KSMKIMVNAALQOFHSPFFREEDPSRQIEPSIQEDHQVPIW 1143  
DB 1072 RDVSELPILGITSFNKSLYEGINAENHRT-KITVIFLKEBETSRSLP----- 1116  
QY 1144 IIVGSLGGLALLLALLLALRKLGFRRSARRRREGLDPT 1183  
DB 1117 LIIGSSIGLLVVLVITAILFKCGFFK--RKYQQLNEST 1154

RESULT 15  
ID ITAE HUMAN STANDARD; PRT; 1179 AA.  
AC P38570: Q9NZU9.  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-E precursor (Mucosal lymphocyte-1 antigen) (HML-1  
DE antigen) (CD103 antigen) (Integrin alpha-TEL).  
GN ITGAE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A., AND SEQUENCE OF 19-38 AND 179-188.  
RC TISSUE=Leukemia, and Lymphocytes;  
RX MEDLINE=94164962; PubMed=8119947;  
RA Shaw S.K., Cepak K.L., Murphy E.A., Russell G.J., Brenner M.B.,  
RA Parker C.M.;  
RT "Molecular cloning of the human mucosal lymphocyte integrin alpha E  
RT subunit. Unusual structure and restricted RNA distribution.";  
RL J. Biol. Chem. 269:6016-6025(1994).  
RN [2]  
RN REVISIONS TO 88-114.  
RP Parker C.M.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE OF 53-1179 FROM N.A.  
RC TISSUE=Fetal kidney;  
RX MEDLINE=20138496; PubMed=10673275;  
RA Touchman J.W., Anikster Y., Dietrich N.L., Maduro V.V., McDowell G.,  
RA Shotelesuk V., Bouffard G.G., Beckstrom-Sternberg S.M., Gahl W.A.,  
RA Green E.D.;  
RT "The genomic region encompassing the nephropathic cystinosis gene  
RT (CTNG): complete sequencing of a 200-kb segment and discovery of a  
RT novel gene within the common cystinosis-causing deletion.";  
RL Genome Res. 10:165-173(2000).  
RN [4]  
RN MUTAGENESIS OF ASP-109 AND PHE-316.  
RX MEDLINE=20400502; PubMed=10837471;  
RA Higgins J.M.H., Cernadas M., Tan K., Irie A., Wang J.-H., Takada Y.,  
RA Brenner M.B.;  
RT "The role of alpha and beta chains in ligand recognition by beta 7  
RT integrins";  
RL J. Biol. Chem. 275:25652-25664(2000).  
CC -1- FUNCTION: INTEGRIN ALPHA-E/BETA-7 IS A RECEPTOR FOR E-CADHERIN. IT  
CC MEDIATES ADHESION OF INTRA-EPITHELIAL T-LYMPHOCYTES TO EPITHELIAL  
CC CELL MONOLAYERS.  
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA  
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAINS LINKED BY A  
CC DISULFIDE BOND. ALPHA-E ASSOCIATES WITH BETA-7.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: EXPRESSED ON A SUBCLASS OF T-LYMPHOCYTES KNOWN  
CC AS INTRA-EPITHELIAL LYMPHOCYTES WHICH ARE LOCATED BETWEEN MUCOSAL  
CC EPITHELIAL CELLS.  
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -1- SIMILARITY: Contains 1 VWFA domain.  
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD103 entry;  
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd103.htm".  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC  
CC EMBL; L25851; AAB59359.2; --

DR EMBL; AF168787; AAF43107.1; -.  
DR PIR; A53213; A53213.  
DR HSSP; P11215; LABX.  
DR Genew; HGNC:6147; ITGAE.  
DR MiM; 604682; -.  
DR GO; GO:0008305; C: integrin complex; TAS.  
DR GO; GO:0004895; F: cell adhesion receptor activity; TAS.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP 3.  
DR Pfam; PF00357; integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PRO1185; INTEGRINA.  
DR PRINTS; PRO0453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 3.  
DR SMART; SM00327; vwa; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS02334; VWFA; 1.  
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
KW Signal; Repeat; Polymorphism; Magnesium; Calcium.  
FT SIGNAL 1 18  
FT CHAIN 19 1179 INTEGRIN ALPHA-E.  
FT CHAIN 19 1177 INTEGRIN ALPHA-E LIGHT CHAIN.  
FT CHAIN 179 1177 INTEGRIN ALPHA-E HEAVY CHAIN.  
FT DOMAIN 19 1124 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 1125 1147 POTENTIAL.  
FT DOMAIN 1148 1179 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 181 198 GLU-RICH (ACIDIC).  
FT REPEAT ? ? FG-GAP 1.  
FT REPEAT ? ? FG-GAP 2.  
FT REPEAT ? ? X-DOMAIN (EXTRA DOMAIN).  
FT REPEAT ? ? VWFA.  
FT REPEAT 401 456 FG-GAP 3.  
FT REPEAT 457 506 FG-GAP 4.  
FT REPEAT 510 571 FG-GAP 5.  
FT REPEAT 573 638 FG-GAP 6.  
FT REPEAT 641 693 FG-GAP 7.  
FT CA\_BIND 522 530 POTENTIAL.  
FT CA\_BIND 586 594 POTENTIAL.  
FT CA\_BIND 654 662 POTENTIAL.  
FT SITE 1150 1154 GFFKR MOTIF.  
FT DISULFID 70 79 BY SIMILARITY.  
FT DISULFID 126 159 BY SIMILARITY.  
FT DISULFID 706 762 BY SIMILARITY.  
FT DISULFID 823 829 BY SIMILARITY.  
FT DISULFID 833 907 BY SIMILARITY.  
FT DISULFID 1008 1033 BY SIMILARITY.  
FT DISULFID 1041 1057 BY SIMILARITY.  
FT CARBOHYD 49 49 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 271 271 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 321 321 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 444 444 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 726 726 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 782 782 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 857 857 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 934 934 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 954 954 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1065 1065 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1096 1096 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT VARIANT 360 360 D -> E.  
FT VARIANT 1041 1041 /FTID=VAR\_008884.  
FT VARIANT 1041 1041 C -> S.  
FT MUTAGEN 208 208 D->A: LOSS OF E-CADHERIN BINDING.  
FT MUTAGEN 316 316 F->A: LOSS OF E-CADHERIN BINDING.  
FT CONFLICT 477 477 V -> I (IN REF. 3).  
FT CONFLICT 482 482 Q -> R (IN REF. 3).  
FT CONFLICT 950 950 R -> W (IN REF. 3).  
FT CONFLICT 1019 1019 A -> V (IN REF. 3).  
SQ SEQUENCE 1179 AA; 130088 MW; E558902EDF9D95E1 CRC64;  
Query Match 14.9%; Score 924.5; DB 1; Length 1179;  
Best Local Similarity 26.3%; Pred. No. 6.9e-52;

Matches	338;	Conservative	206;	Mismatches	482;	Indels	259;	Gaps	52;
Qy	23	FNMTRKRVIPGSR	TAFFGVTYVQOHD	ISGNK--KLVV	GAPLETNGYQKTGDV	KPCV	VIHG	81	
Db	19	FNDVAPRWLT	PKGGAFLVLSLL	HDPSTNQTLW	TSP---RTKETP	GPLH	CS	75	
Qy	82	NCTKLNLRVY	LSNVSRKDNRL	GLSLATNPKD	NSFLAC	SPLM--SHECG	SSYYTGM	138	
Db	76	EILCHPV	EHVPIPKGRH	-----GVTV	RS--HHGV	LICIQLV	RPHSL	125	
Qy	139	CS						159	
Db	126	CSLLGDLR	PQAQANFFD	LENLDPD	RVDTGDCYSN	KEGGEDVNTAR	QRALEKEE	185	
Qy	160	-----QTYM	DIVILDCG	NSIYP--WVEV	QHFLINIL	KFEYIPGQ	IQGVVQ	205	
Db	186	EDKEE	EDDEBER	AGTEAII	LDGSGDIP	DFQAKDFISNM	KNFYECFC	245	
Qy	206	YGEDV	HEFLNDRSV	KDVVEA	ASHTEQRG	TETRTAF	GIEPAR	263	
Db	246	YGGVIQ	TEFLDRSD	QVNASLAR	VONITQ	GVSV--TKTAS	AQHVLD	304	
Qy	264	KWIMIV	DGESH	DSP--DLEK	VIQOES	EDNTRV	AVAVL	317	
Db	305	KWVVL	TGGGIF	EDPLNLT	TVINS	PKQGV	VERPAIGV	354	
Qy	318	KYTASD	DDKXHF	NVTEA	AKIYD	ALCDBRI	SLEGTN	377	
Db	355	NLTASD	PDETH	AFKVTY	NALDGL	LKLRVNI	ISMEGT	412	
Qy	378	D--GVL	LGA	GAYDN--GAVL	KETSAG	KVIFLR	BSYLKE	435	
Db	413	ERVLL	LGA	VAFD	WSGG	ALLYDTR	SRGRFL	470	
Qy	436	RQEVY	YAGAPR	FNHTK	VLFT	MNNRS	TIHC	495	
Db	471	TCSLSY	YAGAPQYKH	GA	VFEL--QKE	GRASFL	PVLEG	529	
Qy	496	DVLV	GAPMYF	NEGR	ERK	VYVYEL--RQR	FVYNG	553	
Db	530	DFLL	VAAPFV	HHVH	G--EBGR	VVYVY	RLS	588	
Qy	554	QDSY	NVYV	GA	LE-----DN	HA--G	AYI	607	
Db	589	QDKL	UVAIG	APLE	GG	ADG	ASFG	648	
Qy	608	IHGOLD	LNED	GLD	LAV	GALGN	AVIL	667	
Db	649	MAGG	FDIS	GLAD	ITV	TGLQ	AVP	699	
Qy	668	CLAA	FCTP	IFL	APHP	QTTV	GR---Y	710	
Db	700	VWV	RLCFE---	ISSV	TAS	SGUR	EALL	756	
Qy	711	GGDR	FTNR	AVLL	SSG	QELC	ERIN	765	
Db	757	SGS	Q	LED	LLMPT	TE	GE	807	
Qy	766	GWPT	L	RVSP	FWP	FWNG	CNE	825	
Db	808	YTP	PAIF	Q	LPY	EK	AKN	842	
Qy	826	TTVP	II	EST	RQ	RAVE	ATLEN	885	
Db	843	E--LV	V	GL	KELT	L	N	897	
Qy	886	RLQ	QKV--CNV	S	YPFR--AK	A	V	940	
Db	898	PVAS	IL	MC	RI	GH	PV	953	
Qy	941	EDN	V	AP	L	R	FL	994	
Db	954	NE--THT	Q	PR	GH	F	V	998	

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 17:53:06 ; Search time 57 seconds  
(without alignments)  
6576.064 Million cell updates/sec

Title: US-09-980-403-2  
Perfect score: 6224  
Sequence: 1 MDLPRGLVVAWALSMPGFT.....PRSARRRREPCLDTPKVLK 1188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL\_25:.\*  
1: sp\_archaea:.\*  
2: sp\_bacteria:.\*  
3: sp\_fungi:.\*  
4: sp\_human:.\*  
5: sp\_invertebrate:.\*  
6: sp\_mammal:.\*  
7: sp\_mhc:.\*  
8: sp\_organelle:.\*  
9: sp\_phase:.\*  
10: sp\_plant:.\*  
11: sp\_rhodent:.\*  
12: sp\_virus:.\*  
13: sp\_vertebrate:.\*  
14: sp\_unclassified:.\*  
15: sp\_rvirus:.\*  
16: sp\_bacteriap:.\*  
17: sp\_archaeap:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5690	91.4	1188	11 Q7TQC3	Q7TQC3 mus musculus
2	4272	68.6	823	4 Q8WY18	Q8WY18 homo sapien
3	3840	61.7	823	11 Q8CB84	Q8CB84 mus musculus
4	2135.5	34.3	1171	13 Q42094	Q42094 gallus galli
5	1196	19.2	1161	11 Q8QVE7	Q8QVE7 rattus norv
6	1160	18.6	1169	11 Q9QXH4	Q9QXH4 mus musculus
7	1131.5	18.2	1169	4 Q8IV46	Q8IV46 homo sapien
8	1125.5	18.1	1196	13 Q98TF1	Q98TF1 cyprinus ca
9	1096.5	17.6	1151	11 Q9JT30	Q9JT30 rattus norv
10	1065.5	17.1	1161	11 Q9WTV4	Q9WTV4 mus musculus
11	1063	17.1	1160	11 Q9R200	Q9R200 mus musculus
12	1045	16.8	1187	13 Q98TF0	Q98TF0 cyprinus ca
13	1040	16.7	780	13 Q06271	Q06271 xenopus lae
14	1018.5	16.4	1086	4 Q96HBI	Q96HBI homo sapien
15	979	15.7	1167	11 Q88341	Q88341 rattus norv
16	954	15.3	1167	11 Q88340	Q88340 rattus norv

17 904.5 14.5 927 6 Q8HZV0  
18 900 14.5 1160 6 Q8MKP4  
19 895 14.4 920 6 Q28984  
20 891 14.3 1038 11 Q8BS01  
21 881 14.2 1332 5 Q9BP08  
22 856 13.8 895 11 Q9WUF8  
23 848.5 13.8 1036 11 Q8CA73  
24 795.5 12.8 1036 11 Q91YDS  
25 742 11.9 1041 5 Q9UB90  
26 738 11.9 1041 5 Q76378  
27 719.5 11.6 1054 5 Q9UGS1  
28 691.5 11.1 1032 11 Q61989  
29 691.5 11.1 1033 6 Q9BGU3  
30 667.5 10.7 323 11 Q8CB84  
31 630 10.1 1119 5 Q86888  
32 604.5 9.7 1034 13 Q98TT7  
33 597 9.6 632 4 Q8NGH6  
34 595.5 9.6 1053 11 Q8OYPS  
35 587 9.4 1132 11 Q80218  
36 584 9.4 1033 13 Q42598  
37 582.5 9.4 1073 11 Q8CC06  
38 571.5 9.2 1474 5 Q86887  
39 567.5 9.1 1016 13 Q91779  
40 556 8.9 604 11 Q8BQ25  
41 551.5 8.9 974 11 Q924W2  
42 547 8.8 1049 5 Q8SY51  
43 545 8.8 1047 6 Q9MZD6  
44 544.5 8.7 1007 6 Q9GK48  
45 539.5 8.7 833 5 Q9BFQ7

#### ALIGNMENTS

#### RESULT 1

Q7TQC3 ID Q7TQC3 PRELIMINARY; PRT; 1188 AA.  
AC Q7TQC3;  
DT 01-OCT-2003 (T-EMBLrel. 25, Created)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
DE All integrin.  
GN ITGAL1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Tiger C.-F., Popova S.N., Velling T., Ekman S., Forsberg E.,  
RA Gullberg D.;  
RT "allbi integrin is important for mesenchymal cell function:  
RT elimination of allbi leads to dwarfism."  
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Johansson M., Popova S.N.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY124460; AAM62130.1; -;  
KW Integrin.  
SQ SEQUENCE 1188 AA; 133067 MW; 69B2BF20D75E4271 CRC64;

Query Match 91.4%; Score 5690; DB 11; Length 1188;  
Best Local Similarity 90.2%; Pred. No. 0;  
Matches 1071; Conservative 57; Mismatches 60; Indels 0; Gaps 0;  
Qy 1 MDLPRGLVVAWALSMPGFTDTFNMDTRKPRVTPGSKTAPFGYTVQOHDISGNKWLVA 60  
Db 1 MDLPRGLVVAWALSMPGFTDTFNMDTRKPRVTPGSKTAPFGYTVQOHDISGNKWLVA 60  
Qy 61 PLFTNGYQKTDYKCPVHGNCTKLNGLRVTLNVSVSRKNNRRLGLSLATNPKDNSFLA 120  
Db 61 PMETNGHQTKGVDYKCPVHGNCTKLNGLRVTLNVSVSRKNNRRLGLSLATNPKDNSFLA 120

QY 121 CSPLMSECCSSYTTTGMCSRVNSNFRFSKTVAPALORCOTYNDIIVILDGNSIYPWVE 180  
 DB 121 CSPLMSECCSSYTTTGMCSRVNSNFRFSKTVAPALORCOTYNDIIVILDGNSIYPWVE 180  
 QY 181 VOHFLNILKFFIIGPQIQGVVQYCEVDVHFFHLNDYRSVXDVVAAASHIEQRGSTET 240  
 DB 181 VOHFLNILKFFIIGPQIQGVVQYCEVDVHFFHLNDYRSVXDVVAAASHIEQRGSTET 240  
 QY 241 RTAFGIEFARSEAFQGRGAKKXNVIITDGHSDSPOLEKVIQSEEDNVTYAVAVL 300  
 DB 241 RTAFGIEFARSEAFQGRGAKKXNVIITDGHSDSPOLEKVIQSEEDNVTYAVAVL 300  
 QY 301 GYNNRGINPETELNEIKYIASDDKGFENVVDEAALKDIDVADGDRIFSLEGTKNET 360  
 DB 301 GYNNRGINPETELNEIKYIASDDKGFENVVDEAALKDIDVADGDRIFSLEGTKNET 360  
 QY 361 SFGLEMSOTGFSSHVEDGVLGAVGAYDNGAVLTKETSAGKVIPLRESYLKFFPELKN 420  
 DB 361 SFGLEMSOTGFSSHVEDGVLGAVGAYDNGAVLTKETSAGKVIPLRESYLKFFPELKN 420  
 QY 421 HGAYLGYTIVTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSLSLTHQMRGQOIGSYF 480  
 DB 421 HGAYLGYTIVTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSLSLTHQMRGQOIGSYF 480  
 QY 481 GSEITSVDIDGCVTDVLLVGAPMYPNEGRERKQVYVYELQNRQFVYNGTLKDSHSYQNA 540  
 DB 481 GSEITSVDIDGCVTDVLLVGAPMYPNEGRERKQVYVYELQNRQFVYNGTLKDSHSYQNA 540  
 QY 541 RFGSSIASVRDLNQDSVNDVVGAPLEDNHAGAIYIFHGFPGSILKTPKORITASELATG 600  
 DB 541 RFGSSIASVRDLNQDSVNDVVGAPLEDNHAGAIYIFHGFPGSILKTPKORITASELATG 600  
 QY 601 LQVFGSIHQGLDNLNEDGLDLAVGALGNVILWSRPVVOINASLHFPESKINIPIHRCCK 660  
 DB 601 LQVFGSIHQGLDNLNEDGLDLAVGALGNVILWSRPVVOINASLHFPESKINIPIHRCCK 660  
 QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYPTRAHLDEGGDRFTRAV 720  
 DB 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYPTRAHLDEGGDRFTRAV 720  
 QY 721 LLSGGQELCERINPHVLDADYVKPVTFVSVEYSLDPPDHGPMDDGMPITLRSVFPFWNG 780  
 DB 721 LLSGGQELCERINPHVLDADYVKPVTFVSVEYSLDPPDHGPMDDGMPITLRSVFPFWNG 780  
 QY 781 CNEDEHCVPLVLDARSDLTAMEYCORVLRKPAQDCSAYTSLSPDTTTFVITESTRORAV 840  
 DB 781 CNEDEHCVPLVLDARSDLTAMEYCORVLRKPAQDCSAYTSLSPDTTTFVITESTRORAV 840  
 QY 841 EATLENREGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900  
 DB 841 EATLENREGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900  
 QY 901 RAKAKVAFRLDSEFSKSIPLHHLIELHLAGSDSNEDSTKEDNVAPLPHLKYEADVLPT 960  
 DB 901 RAKAKVAFRLDSEFSKSIPLHHLIELHLAGSDSNEDSTKEDNVAPLPHLKYEADVLPT 960  
 QY 961 RSSLSLHYEVLKANSLSERYDGIPTPFCIPRIQNLGLPPIHGMKMTTPIATRSNRL 1020  
 DB 961 RSSLSLHYEVLKANSLSERYDGIPTPFCIPRIQNLGLPPIHGMKMTTPIATRSNRL 1020  
 QY 1021 KLADFTLDEANTSCNIGNSTERYPTPVEEDLRRAPQLNHNSDVVSNINRLVNPQEI 1080  
 DB 1021 KLADFTLDEANTSCNIGNSTERYPTPVEEDLRRAPQLNHNSDVVSNINRLVNPQEI 1080  
 QY 1081 NEHLGLNLWLSLKALYKSMKIMVNAALQOFHSPPIFEEDPSPQIPEIISKQEDWQV 1140  
 DB 1081 NEHLGLNLWLSLKALYKSMKIMVNAALQOFHSPPIFEEDPSPQIPEIISKQEDWQV 1140  
 QY 1141 PIWIVGSTLGGLLALLVALLKLGFFRSARRRPPGLDPTPKVLE 1188  
 DB 1141 PIWIVGSTLGGLLALLVALLKLGFFRSARRRPPGLDPTPKVLE 1188

RESULT 2  
 Q8WYI8 PRELIMINARY; PRT; 823 AA.  
 ID Q8WYI8;  
 AC Q8WYI8;  
 DT 01-MAR-2002 (TREMELrel. 20, Created)  
 DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
 DE MSTP018.  
 GN MST018.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Aorta;  
 RA Xu Y.Y., Sun L.Z., Wu Q.Y., Liu Y.Q., Liu B., Zhao B., Wang X.Y.,  
 Song L., Ye J., Sheng H., Gao Y., Zhang C.L., Zhang J., Wei Y.J.,  
 Sun Y.H., Jiang Y.X., Zhao X.W., Liu S., Liu L.S., Ding J.P.,  
 Gao R.L., Qiang B.Q., Yuan J.G., Liew C.C., Zhao M.S., Rui R.T.;  
 RA Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AF111799; AAL39001.1;  
 DR GO; GO:0008305; C:integrin complex; IEA.  
 DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.  
 DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR PRINTS; PRO1185; INTEGRINA.  
 DR SMART; SM00191; Int\_alpha; 4.  
 SQ SEQUENCE 823 AA; 92672 MW; DE4E78079DCD4925 CRC64;

Query Match 68.6%; Score 4272; DB 4; length 823;  
 Best Local Similarity 99.6%; Pred. No. 2.9e-288;  
 Matches 820; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 366 MSOTGFSSHVEDGVLGAVGAYDNGAVLTKETSAGKVIPLRESYLKFFPELKNHAYL 425  
 DB 1 MSOTGFSSHVEDGVLGAVGAYDNGAVLTKETSAGKVIPLRESYLKFFPELKNHAYL 60  
 QY 426 GYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSLSLTHQMRGQOIGSYFSGEIT 485  
 DB 61 GYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSLSLTHQMRGQOIGSYFSGEIT 120  
 QY 486 SYVDIDGCVTDVLLVGAPMYPNEGRERKQVYVYELQNRQFVYNGTLKDSHSYQNAFPSS 545  
 DB 121 SYVDIDGCVTDVLLVGAPMYPNEGRERKQVYVYELQNRQFVYNGTLKDSHSYQNAFPSS 180  
 QY 546 TASVRDLNQDSVNDVVGAPLEDNHAGAIYIFHGFPGSILKTPKORITASELATGLQYFG 605  
 DB 181 TASVRDLNQDSVNDVVGAPLEDNHAGAIYIFHGFPGSILKTPKORITASELATGLQYFG 240  
 QY 606 CSIHQGLDNLNEDGLDLAVGALGNVILWSRPVVOINASLHFPESKINIPIHRCCKRSGRD 665  
 DB 241 CSIHQGLDNLNEDGLDLAVGALGNVILWSRPVVOINASLHFPESKINIPIHRCCKRSGRD 300  
 QY 666 ATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYPTRAHLDEGGDRFTRAVLLSSG 725  
 DB 301 ATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYPTRAHLDEGGDRFTRAVLLSSG 360  
 QY 726 QELCERINPHVLDADYVKPVTFVSVEYSLDPPDHGPMDDGMPITLRSVFPFWNGCNEDE 785  
 DB 361 QELCERINPHVLDADYVKPVTFVSVEYSLDPPDHGPMDDGMPITLRSVFPFWNGCNEDE 420  
 QY 786 HCVPLVLDARSDLTAMEYCORVLRKPAQDCSAYTSLSPDTTTFVITESTRORAVAEATLE 845  
 DB 421 HCVPLVLDARSDLTAMEYCORVLRKPAQDCSAYTSLSPDTTTFVITESTRORAVAEATLE 480  
 QY 846 NREGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQVCNVSYPFPRAKAK 905  
 DB 481 NREGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQVCNVSYPFPRAKAK 540  
 QY 906 VAFRLDSEFSKSIPLHHLIELHLAGSDSNEDSTKEDNVAPLPHLKYEADVLTRSSSL 965

```

Db 541 VAERLDFEFSKSLFLHLEIELAAGSDSNERDSTKEDNVAPLRFHUKYEDVLFTRSSSL 600
Qy 966 SHYEVKLNLSERYDGIQPPFCIPRIQNLGFPPIHGMWKITIPATISGNRLKLRDF 1025
Db 601 SHYEVKLNLSERYDGIQPPFCIPRIQNLGFPPIHGMWKITIPATISGNRLKLRDF 660
Qy 1026 LTDEANTSNIWGNSTYRPTVEEDLRAPQLNHSNDSVSTNCNIRLVNPOEINFHLL 1085
Db 661 LTDEANTSNIWGNSTYRPTVEEDLRAPQLNHSNDSVSTNCNIRLVNPOEINFHLL 720
Qy 1086 GNLWLSLXALKYKSKMKNVNAALQRFHSPRIFREEDPSRQIETFEISKQEDQWQVPIWII 1145
Db 721 GNLWLSLXALKYKSKMKNVNAALQRFHSPRIFREEDPSRQIETFEISKQEDQWQVPIWII 780
Qy 1146 VGSTLGGLLLLALLVLALRWLKLGFRRSARRRREGLDPTPKVLE 1188
Db 781 VGSTLGGLLLLALLVLALRWLKLGFRRSARRRREGLDPTPKVLE 823

RESULT 3
Q8CE84 PRELIMINARY; PRT; 823 AA.
AC Q8CE84;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE MSTP018 homolog.
GN 473245H24RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium.
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK028821; BAC26137.1; -.
DR MGD; MGI:2442114; 473245H24RIK.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01819; FG-GAP; 3.
DR PRINTS; PR01185; INTEGRINA.
DR SMART; SM00191; Int_alpha; 4.
SQ SEQUENCE 823 AA; 92264 MW; A330236324A08089 CRC64;

Query Match 61.7%; Score 3940; DB 11; Length 823;
Best Local Similarity 87.8%; Pred. No. 3.4e-258;
Matches 723; Conservative 51; Mismatches 49; Indels 0; Gaps 0;

Qy 366 MSQTGSSHHVEDGVLGVAGYDMNGAVLKETSACKVPLRESYLKEPPELKHGAYL 425
Db 1 MSQTGSSHHVEDGILLGVAGYDMNGAVLKETSACKVPHRESYLKEPPELKHAAYL 60
Qy 426 GYTVTSVWSRQGVTVVAGAPFNHTGKVLFTMHNNRSLTIHQAMRGQIGSYFGSEIT 485
Db 61 GYTVTSVWSRQGVTVVAGAPFNHTGKVLFSMNNRSLTIHQALRGQIGSYFGSEIT 120
Qy 486 SVIDIDCGVTDVLLVGPAMYFNEGREGKVVYVLEFQNEFVYNGTLKDSHSHYONARFGSS 545
Db 121 SVDVNDVRVTDVLLVGPAMYFSEGRERGVYVYVNFQNEFVYNGTLKDSHSHYONARFGSC 180
Qy 546 IASVRDLNDSYNDVVVGVAPLENDHAGATYIPHGFRGSLTKPKORITASELATGLOYFG 605
Db 181 IASVQDLNQSDNDVVVGVAPLEDSHRGATYIPHGFTNLKPKORITASELATGLOHFG 240

```

```

Qy 606 CSIHQOLDNLDELIDLAVALGNVILWSRPVQINASLHFEPSKINFHEDCKSGRD 665
Db 241 CSIHQOLDNLDELIDLAVALGNVILWSRPVQINASLHFEPSKINFHEDCKSGRD 300
Qy 666 ATCLAAFICFTPIFLAPHFQTTTGVIRYNATMDERRYTPRAHLDEGGRFTNRAVLLSSG 725
Db 301 ATCLAAFICFTPIFLAPHFQTTTGVIRYNATMDERRYTPRAHLDEGGRFTNRAVLLSSG 360
Qy 726 QELCERINPHVLDIADYKVPVTFSEVSLDEPDHGMDDGCHPTTLRVSVFPMWNCNEDE 785
Db 361 QELCERINPHVLDIADYKVPVTFSEVSLDEPDHGMDDGCHPTTLRVSVFPMWNCNEDE 420
Qy 786 HCVPLDLVLDARSDLPTAMEYCORVLKPAQDCSAYTSLFDTVTTFIESTRORVAVEATLE 845
Db 421 HCVPLDLVLDARSDLPTAMEYCORVLKPAQDCSAYTSLFDTVTTFIESTRORVAVEATLE 480
Qy 846 NRGENAYSTVLNISOSANLQFASLIQKEDSDGSECVNEBRRLQKQVNCVSPFPRAKAK 905
Db 481 NRGENAYSAVLNISQSENLOQFASLIQKEDSDGSECVNEBRRLHKKVNCVSPFPRAKAK 540
Qy 906 VAFRLDSEFSKSIPLHHLIEIELAAGSDSNERDSTKEDNVAPLRFHUKYEDVLFTRSSSL 965
Db 541 VAFRLDSEFSKSVFLHHLQIHLGAGSDSHQEDSTADNTALLRFHUKYEDVLFTRSSSL 600
Qy 966 SHYEVKLNLSERYDGIQPPFCIPRIQNLGFPPIHGMWKITIPATISGNRLKLRDF 1025
Db 601 SHYEVKLNLSERYDGIQPPFCIPRIQNLGFPPIHGMWKITIPATISGNRLKLRDF 660
Qy 1026 LTDEANTSNIWGNSTYRPTVEEDLRAPQLNHSNDSVSTNCNIRLVNPOEINFHLL 1085
Db 661 LTDEANTSNIWGNSTYRPTVEEDLRAPQLNHSNDSVSTNCNIRLVNPOEINFHLL 720
Qy 1086 GNLWLSLXALKYKSKMKNVNAALQRFHSPRIFREEDPSRQIETFEISKQEDQWQVPIWII 1145
Db 721 GNLWLSLXALKYKSKMKNVNAALQRFHSPRIFREEDPSRQIETFEISKQEDQWQVPIWII 780
Qy 1146 VGSTLGGLLLLALLVLALRWLKLGFRRSARRRREGLDPTPKVLE 1188
Db 781 VGSTLGGLLLLALLVLALRWLKLGFRRSARRRREGLDPTPKVLE 823

RESULT 4
Q842094 PRELIMINARY; PRT; 1171 AA.
AC Q842094;
DT 01-JAN-1998 (TremBLrel. 05, Created)
DT 01-JAN-1998 (TremBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE ALPHAI integrin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gizzard;
RX MEDLINE=97476270; PubMed=9334246;
RA Obata H., Hayashi K., Nishida W., Moniyama T., Uchida A., Ochi T.,
RA Sobue K.;
RT "Smooth muscle cell phenotype-dependent transcriptional regulation of
RT -the alpha1 integrin gene.";
RL J. Biol. Chem. 272:26643-26651(1997).
DR EMBL; AB000470; BAA23160.1; -.
DR EMBL; AB000471; BAA23161.1; -.
DR HSSP; P17301; IAOX.
DR PIR; A55348; A55348.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.

```



Db	8	LLCGWVILASCHG-----SNLDVEBP-IVFREDAASFGQTVVQF-----GGSLVVGAPLEAVA 59
Qy	67	YQKGTGVYKCPVTHGNCITKMLGRVTLNSVSEKDNKRLGLSLATMPKONSLFACSPWLS 126
Db	60	VNQTGRLYDCAPATGMCQPIVL--RSPLEAV-----NMSLGLSLVATATNNAQLLACGPTAQ 113
Qy	127	HECGSSYVTTGMCSSRVNSNFRPSKTVAPALQRC-QTYMDIVIVLQGSNISI--YPWVEVOH 183
Db	114	RACVKNMYAGKSCLLIGSSLOFTQOAVPASMPECPQEMDIAFLIDGSGINORDFAQMKD 173
Qy	184	FLINILKKFYIGPGQIQGVGVQGEDVHSHFHLNDYRSKDVVVEAAASHIEQOQGTETRTA 243
Db	174	EVKALMGEF--ASTSTLPSLAQYSNILKTHFTTEPKNILDQSLVDPIVLQQL--TYTA 230
Qy	244	FGIEFARSEAF--QKGRGKAGKMWIVITDGBHSDP-DLEKVIQOSEDNDVTRVAVAVL 300
Db	231	TGTRTMEELFHSKNGSRSAKKILLIVTDQKYRDPLEYSQVIPAADKAGIIRYAIQVG 290
Qy	301	GYNNRRGINDETFLNEIKYIASDPDDKHFFNVVDEAALKDVIDALGDRFSLGRT-NKNE 359
Db	291	DAFOE-----PTALKELNTIGSAPPQDHVFVKGNFPAALRSIQLOLEKFAIECTGVSRS 345
Qy	360	TSFGLEMSQTFSSHVVVDCGLVGAQVADNNGAVILKETSAGKVP--LRESVLEKFPBE 417
Db	346	SSPOHEMSQGFSSALTDGCPVLGAVGFSFNSGGAF-----LYPNTRTPTINMSQEN 398
Qy	418	LKNHGAYLGYTVTSVYSSROGRVYVAGAPENHTGKVLFTMENNELSLTHQAMRCQQIG 477
Db	399	VDMRDSILGYS-TAVAFWKGVHSLILGAPRHQHTGKVVIPT-QEAKRWPKSEVRQTQIG 456
Qy	478	SYFGESEITSDIDGDGVTDLVLLGAPMYFNECREGRKVVYYELR--QNRPVYNGTLKDSH 535
Db	457	SYFGASLCSVDVDRDGGSTDLVLGAPHYYEQTR--GGQVSVPFPGVGRGAQCEATLHGEQ 515
Qy	536	SYQNAERGSSIASVRDLNQSYNDVVVGAPLENDHAGAIYIFPG-FRGSILKTPKORITA 594
Db	516	GHPWGRFGVALTVLGVDVNGDNLADVAIGAPGEESRGAVYIFHGASLEIMPSPSORVTG 575
Qy	595	SELATGLQYGCSTHGGQLDNEDQLIDLAVGALGNVILMSRPVVQINASLHFEPSKINI 654
Db	576	SQSLRLQYFGQSLSGGQDITQDGLVDLAVGAQHVLRLSLPLLKVELSIRFAPMEVAK 635
Qy	655	FHRDKRSGRATC-----AAPLCPTFLFAPHQTTTVG-IRKNATMDEERYTPRAHLDE 710
Db	636	AVYQCMF--RTPTVLEAGEATVCLTHKGSPDILLGNVQSVRYDLALDQRLISRAIFDE 693
Qy	711	GGDRFTNRAVLSSGQELCRLINPHVLID-TADVYKPTFSVEYSL-----EDPDHGPMLD 764
Db	694	TKNCTLTRKTYLGLGDH-CETVKLLIJDVEDAVSPILRLNLSLVRDSSAPENLHPVLA 752
Qy	765	DGMPTTLRVSPVFMNCGNEDEHCVPDILVDLARSDLPTAMBYCYORVLKRPQDCSAYTLSP 824
Db	753	VGSQDHITASLPFPEKNCKQELLCEGL-----GISENF 785
Qy	825	DTTVFIIETQRVAVAEATLENRGENAYTVMLISQSANLQFASLI---QKEDSDGSIEC 881
Db	786	SGLOVLVVGSGPELTVTVTVNMBEGESYGLTVKFPYPAGISYRRTVGTQPHQVPIRLAC 845
Qy	882	VNE---ERRLOQVCNVSYPFFRAKAKAVAFRLDSEPSKIFL-HHLEIELAAGSDSNERD 937
Db	846	EAEPAAQEDLRSSCSINHPIFRGAKTTFMIITDVSYKAFGLGDLRLAKAKASSENKPD 905
Qy	938	STKEDNVAPLRFHLKYRADVLFTRSSLSHYEVKLNLSRLRYDYGIGPPFPSCIPIRIQMLGL 997
Db	906	TNK--TAFQLELPVKYTVYTLISRQEDSTNH-VNPFSS--HGGRRQEAHRYRVNMLSP 959
Qy	998	PPHGMWMMKIPIATIRSGNRLKLRLDFTDEANTSCNIGNSTEYETPTVEE-----DL 1052
Db	960	LKL-AVRNVFWPVL-----LNGVAVMDVTLSPPAQGVSCVSKPKQPDFTQI 1009
Qy	1053	RRAPQLNHSNDVVSINCNIRLVPNQ-EINPHLLIGNL---WLBSL---KALKYKSKMIMV 1105
Db	1010	ORRVLDCSIADCLHFRCDIPSLDIOBELDPIPLNGLSFGWVSOTLQEKVLLVSEARITF 1069

```

Qy 1106 NAALQRFQSPFTFREDDPSRQIEFISKOEDWV--PWIIVGWSVLGGLLLLALIVLAL 1166
Db 1070 DTSVYQLPQGAFLR---AQVETTL---EYVVYEPFLVAGSSVGGLLLLALITVVL 1122
Qy 1164 RKLGFER 1170
Db 1123 YKLGPFK 1129

RESULT 6
ID Q9QXH4 PRELIMINARY; PRT; 1169 AA.
AC Q9QXH4;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Leukocyte adhesion glycoprotein p150,95 alpha integrin subunit.
GN ITGAX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
XP [1]
RP SEQUENCE FROM N.A.
RA Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
RA Teuchiya H.;
RT *Isolation of Genes Selectively Expressed by Dendritic Cells.*;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF111864; AAP23492.1; --
DR HSPB; P11215; IBHQ
DR MGD; MGI:196509; Itgax.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004995; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP_3_A.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PSS0234; VWFA; 1.
KW Integrin.
SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;

Query Match 18.6%; Score 1160; DB 11; Length 1169;
Best Local Similarity 29.5%; Pred. No. 2,3e-71;
Matches 372; Conservative 215; Mismatches 498; Indels 174; Gaps 49;

Qy 9 VAMALSLWPGFTDT--FNMDTRKPRVIPGSRATF-----FGYTVQOHDISGKNKLWVGA 60
Db 6 IAFLLLI--GFVSLCLGNLDAEK-----LTHFHMDDGAEGHSGVLQYDSS---VWVWGA 53

Qy 61 PLETNGYQKTDGVVVKCPVIEGNTCKLMGLRVTLNSVSEKDNMRGLGLSLATNPKDNSFLA 120
Db 54 PKELKATNQIGGLYKCGYHTGNCEPISL-----QVPPBAVNI-SLGLSAAATNPMSLLA 107

Qy 121 CSPLWSHECCGSYYTTGNCGRVSNRFFSKTVAPALQRC-QTYMDIVIVLDGNSI--YP 177
Db 108 CGPTVHTCRENIYLTGLCFLLSSFFQSQNFTTAQCECPKQODIVFLIDGGSGISSTD 167

Qy 178 WVEVQHFLINILKIFYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDVAEASHIBQGG 237
Db 168 PEKMLDFKAVMSQ--LQRSTRESLNQFSDYFRVHTFNFFNSTSPSLSLGSRQLRG 225

Qy 238 TETRTATGIEIFARSEAF--QKGRKGAKNMVIVITDGESH-DSPDLEKVICQSEDNVTR 294
Db 226 -YTYTSAIKHIVITELPTTSGARQADATKVLIVITDGRKGDNIUSDSVIMAEASIIIR 284

```



DE Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RA Strausberg R.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC038237; AAH38237.1; .  
 DR GO; GO:0008305; C:integrin complex; IEA.  
 DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.  
 DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR02035; VWF\_A.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00357; Integrin\_A; 1.  
 DR PRINTS; PR00092; vwa; 1.  
 DR PRINTS; PR01185; INTEGRIN.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00327; vwa; 1.  
 DR PROSITE; PS00242; INTEGRIN ALPHA; 1.  
 DR PROSITE; PS00234; VWFA; 1.  
 SQ SEQUENCE 1169 AA; 128521 MW; 178484FEFC79BB6 CRC64;  
 Query Match 18.2%; Score 1131.5; DB 4; Length 1169;  
 Best Local Similarity 28.4%; Pred No. 2.2e-69;  
 Matches 33%; Conservative 218; Mismatches 472; Indels 201; Gaps 48;  
 QY 7 LVAVALSILWPGFTDTFNMTRKPRVPGSRTPA?-----PGYVQOHDISGNKWLUVGA 60  
 DB 10 LPTALATSL-----GFNLDTEE-----LTAFRVDSAGFGDSVVQY---ANSWVVGVA 53  
 QY 61 PLETNGYQKTDVYKCPVHIGNCTKLNIGRVTLNVSERKDMRGLGLSLATNPKNSFLA 120  
 DB 54 POKITAAQTGLYQCGYSTGACEFGL-----QVPEAVNMSLGLSLATSTSPQLLA 107  
 QY 121 CSPLMSHEGSSSYTTTGMCSRVNSFRSKTVAPALQRC-QTYMDIVIVLDGNSNLYP-- 177  
 DB 108 CGPTVHECGRMNLTGLCFLGPT-QLTQRLPVSRQECPRCEQDIVFLIDGSGSISRN 166  
 QY 178 WVEVOHFLINILKXYIGPGQIQGVQVQYGVGDVWVEHFLNDYRSKDVVEAASHLEQ-RG 236  
 DB 167 FATMNFVRAVISQFQ--RPSTQFLMDFSNKFTFTHFFRFRSNNPUSLASVHQLQG 224  
 QY 237 GTETRTAFIEFARSEAFQKGRGKAKVMIVITDGESH-DSPDLEKVIQOQSERDNVTRY 295  
 DB 225 FVTATAIQNVHRLPHASYGARRDAAKILVITDGKESLDYKDVIPMAAAGIIRY 284  
 QY 296 AVAV-LGYNRRGINPFTLNEIKVIASDPDKHFNVTDEALDKDIVDALGDRIFSLG 354  
 DB 285 AIGVGLAFQNRNS-----WKELNDIASKPSQEHFKVDFDALKDIQQLKEKIPAIEG 338  
 QY 355 T-NKNETSPGLEMSOTGFSHVVDGLGAVGYDMNGAVLKETSAGKVIP--LRRESYL 411  
 DB 339 TETTSSESELEMAQEGFSAVETPDGVLGAVGSTWAGAF-----LYPPNMSPTFI 391  
 QY 412 KSPPELKNHAYLGTYVTSSVSGRQVYVAGAPRPNHTGKVLFTMNNRSLTIHQAM 471  
 DB 392 NMSQENVDNRDSYLGYS-TELALWKGVSLSLVGAPRYQHTKAVIFT-QVSRQWRMAEV 449  
 QY 472 RQOIGSYSGSRIITSVDIDGCVTDVLLVAGAPMYNEGRGKVVYVLEQ--NRPVNG 529  
 DB 450 TGTQIGSYSGSLCSVDVDSGSTDLVLGAGPHYEQTR-GQVSVCPPLPRGWRWNCDA 508  
 QY 530 TLKDSHSYQNAFSGSIASVRDLNDSYNDVWVVGAPLEDNHAAGIYIFHGFTG-SILKTP 588  
 DB 509 VLYGEQGHGWGFRGAALTVLGVDVNGDKLTDVVGAPGKKNRGAIVLPHGLGSPISPSH 568  
 QY 589 KQRIITASELATGLQYFGCSIHQGLDNEGLDLAVAGNAVILWSRPVVOINASLHFE 648  
 DB 569 SQRIAGSLSRLOVFGQALSGGQDLTQGLVDLAVGARGQVLLLRTRPVLWVGVSQMFI 628

QY 295 YAVAV-LGYNNRRGINPETFLNEIKVIASDPDDKHFFNVTDDEALKDIVDALGDRIFSL 353  
 DB 285 YAVGVGAFYN-----EHSQELKALASPSHSEYVFSVENFDALXDIENQLEKIPAI 338  
 QY 354 GT-NKNETSPGLEMSOTGFSHVVDGLGAVGYDMNGAVLKETSAGKVIPLRSYLK 412  
 DB 339 GTETPSSSTFELMSQGSFSAVETPDGVLGAVGSTWAGAFIYP-----NRPPTFIN 393  
 QY 413 EPPELKNHAYLGTYVTSSVSGRQVYVAGAPRPNHTGKVLFTMNNRSLTIHQAMR 472  
 DB 394 MSQENEDMRDAYLGS-TALAFWKGVHSLILGAPRHQHTGKVIET-QESRHRPKSEVR 451  
 QY 473 GOQIGSYSGSEITSVDIDGCVTDVLLVAGAPMYNEGRGKVVYVLEL--QNRPVYNGT 530  
 DB 452 GTQIGSYSGSLCSVDMDRSGSTDLVLGPHYEHYTR-GQVSVCPMPGVGRWHCGTT 510  
 QY 531 LKDSHSYQNAFSGSIASVRDLNDSYNDVWVVGAPLEDNHAAGIYIFEG-FRGSILKTPK 589  
 DB 511 LKGGQGHGWGFRGAALTVLGVDVNGDKLTDVVGAPGKKNRGAIVLPHGLGSPISPS 570  
 QY 590 ORITASELATGLQYFGCSIHQGLDNEGLDLAVAGNAVILWSRPVVOINASLHFE 649  
 DB 571 QRISASQIPRQIFGOSLSGGQDLTREGLDVAVGSKGRVLLTRTPILRVSPVHTFP 630  
 QY 650 SKINIFHRDCKRS-GRDATCLAAFLCPTPIPLAFPHFOTTV--GIRYNATMDERRVTP 706  
 DB 631 AEISRSYFEGEQVAPETQLSDATVCL-HIESPTQLGDLRSTVTFDLALDHGRILSTRA 689  
 QY 707 HLDGGRDFTNRVALLSGQELCERINPHV-LDTADYVKVPVTSVEVSLD-----PDH 759  
 DB 690 IFKETKTRALTRVTLGLNKH-CESVKLLPACVEDSVTITLRLNPSLVGVPISLQNL 748  
 QY 760 GPMLLDCKPTTLRVSPFPMGNEDEHCVPL-VLDARSPLTAMEYCORVLAKPAQDCS 818  
 DB 749 QPMLAVDDQYFTASLTFEKNCGADHICQDDLVSVPFGPDLKT-----791  
 QY 819 AYTLSDFTTPIIESTRQVAVATELNRGENAYSTVLNISQSANLQF-----ASLIQ 871  
 DB 792 -----LVVGSDELMVDVTVSNDGDSYGTVTILFPVGLSFRRAVAGQVFLRK 840  
 QY 872 KEDSDGS-----IECVNEERLQ--KQVNVSYPPFERAKAVARLDSFS-KSIF 919  
 DB 841 KEDQWQRRGHSLHLMCDSTPDRSQGLMSTCSRRHVPFRGSGMTFLTFVDSRAEL 900  
 QY 920 LHLEIELAAGSDNERDSTKDNVAPLRFHLKYEADVLFTRSSLSHYEVKLNSSLERY 979  
 DB 901 GDRLLRARVGSENNVPQTK--TTFQLELPVKY--AVVTMISSHDQFTKYNLFSTSEK 955  
 QY 980 DGIGPPPSCI---PRIONGLFPHGMMKITPIATRSGRNLLKLADFTDZANISCN 1036  
 DB 956 E---KTSVVEHFRQVANNLQRDV-PVSINFVPIELK-GEAVMTVMVSHPNPLTQC-- 1007  
 QY 1037 WGNSTERYPTVEEDL---RRAPQLNHSNDSVVSINCNIRLVN---QEINFHLLGNL 1088  
 DB 1008 ----YRNLKPTQFDLLTHMQSPVLDCSIADCLHLCDI---PSLGLLDLYFLKGNL 1060  
 QY 1089 ---WLRSL---KALKYKMKIMVNAALQRFHSPFIFRDEDPQRQIFEFISKQEDQV-- 1140  
 DB 1061 SFGMISQTLQKVVLLSEABITFNTSVYSLQPGQAFRAQTKTVLEM-----YKVHN 1113  
 QY 1141 PIWLVGSLGELLALLVLLALRKLGFERSARRRREBGLD-----PTPKVLE 1188  
 DB 1114 EVPLIVSSVGGELLALLIITALLYKAGFP---KQYKEMLEERANGQVSDGTTPQVAQ 1159  
 RESULT 7  
 Q81VAG PRELIMINARY; PRT; 1169 AA.  
 ID Q81VAG  
 AC Q81VAG  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)



```
QY 874 DSDGSECVNEER-RLQKQVQNSVYPPFRAKAKVAF-----RLDSEFSKSTFLHLEI 925
Db 834 RTRSS--CGRDSGATNRTCSIDLVPYRSGTTTQFLGTFRVKNKNDSS-----NRMEI 886
QY 926 ELAAGSDSN-ERSTKEDNVAPLRFHLKYEADVLFTRSSLSHYEVKLSLERYDGIQ 984
Db 887 MITANSNNGNSMSTVERSVQFAV--DLAISLVAEDSVTY---MNFSLJ--DRGP 937
QY 985 -PSCIFERIONLGLFPHGMMKTIPIATRSNRLKLKDLDFITDEANTS-----1033
Db 938 KPLNITYKVENSGR---KDLFVSVTLPCQTPHVLTPHTFSMHEVHVSFISSYHQIIM 994
QY 1034 CNTWGNSTERYPTVEEDLAPOLNHSNDVVSINCN-IRLVNPOBINFHILGNLWLRS 1092
Db 995 CLLNKHLFFSP-----ELSAVQVSTGMSLREVECSQFDLNKSSAVHFNLTADARLQN 1049
QY 1093 LKALKYK-----SKIMVNAALQRFHSPFFREDPSR-QIEFEISKQED 1137
Db 1050 VKEYESKSYFPRKUNVFSIAELNNTSLYNQTSSELKY---NPHRSQTEVKV----E 1102
QY 1138 WQVP---IWIIVGSLGGLLLALLVLAERKLGFFRSAR 1173
Db 1103 FVVPSELMLVCTGAVGSGFFLLIILPLILKCGFFKRR 1141

RESULT 9
Q9J130
ID Q9J130 PRELIMINARY; PRT; 1151 AA.
AC Q9J130;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Integrin beta 2 alpha subunit.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RA Fathallah D.M. Sr., Zerria K. Jr.;
RT "Cloning of the rat CD11b cDNA sequence.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; A2368593; AAF81280.1; -
DR HSSP; P1215; 1BHQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP_3_
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA_1; 1.
SQ SEQUENCE 1151 AA; 126943 MW; 87785695D4074CAS CRC64;

Query Match 17.6%; Score 1096.5; DB 11; Length 1151;
Best Local Similarity 28.8%; Pred. No. 5.8e-67;
Matches 357; Conservative 203; Mismatches 494; Indels 185; Gaps 49;

QY 7 LVVAWALSIPGFTDTFMOTRPRVIPGSRTPFGYVVOOHDSIGNKWLAVGAPLETNG 66
Db 5 VLLATVLTLCG-----FNLDTENPMTQENASG-FGOSVIQ---LGTRVVVAAPQEVKA 56
QY 67 YQKTDGVYKCPVHGNCTKLNGLRVLTSNYSERKDNRLGLSLATNPKDNSFLACSP LMS 126
Db 57 VNQTGALYQCDYSNRCDDIPL-----QVPEAVNMSGLSLAAATVTPVQLACGPTVH 110
```

QY 1095 KVEPYTHVNPVLIVGSSVGLVLLALITAGLYKLGFFK 1133

DB

RESULT 10

Q9WTV4

Q9WTV4 PRELIMINARY; PRT: 1161 AA.

AC Q9WTV4;

DT 01-NOV-1999 (TREMELrel. 12, Created)

DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)

DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)

DE Integrin alpha L.

GN ITGAL.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DBA/2J; TISSUE=Spleen;

RA Ma R.Z.; Teuscher C.;

RL Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AF065902; AAD25885.1; --

DR HSSP; P20701; ILPA.

DR MGD; MGI:96606; Itgal.

DR GO; GO:0008305; C:integrin complex; IEA.

DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.

DR GO; GO:0007160; P:cell-matrix adhesion; IEA.

DR InterPro; IPR000413; Integrin\_alpha.

DR InterPro; IPR02035; VWF\_A.

DR Pfam; PF01839; FG-GAP; 3.

DR Pfam; PF00357; Integrin\_A; 1.

DR Pfam; PF00092; vwa; 1.

DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int\_alpha; 5.

DR SMART; SM00327; VWA; 1.

DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.

DR PROSITE; PS0234; VWFA; 1.

SQ SEQUENCE 1161 AA; 128240 MW; 86B102F7B209B431 CRC64;

QY 398 TSAGKVIPIRE-----SYLKFEFFBELKNHGAYLGVTYTSVSSRQGRVYVAGAPRNFHTG 452

DB

QY 372 --AGFLDLREDLQGAIFVGOEPLTSDVRGGYGYTVAWMTSSRSRPLAAGAPRYQVG 429

QY 453 KVILFTMEN-----NRSLTTHQMRGQIGSYFSGSEITSDVIDDGDVTDVLLVGAAPWFN 507

DB

QY 430 QVLLFQAPAEAGRWNQT-----QKTBGTQIGSYFGCELCSDLDQDGEALLIGAPLPPG 485

QY 508 EGRERGKVVYVELRQNRFPVYNGTLKDSHSYNARFGSSIASVRDLNQSYNDVYVVGAPLE 567

DB

QY 486 EQR--GGRVFTYQRRQSLFEMVSELOQSDPGYPLGRFGAIAITALTIDNGDRLTDAVGA 544

QY 568 DNHAGAIYIFHGFRGSLIKTPKQITASELATGLQYFGCSIHGQDLNEDGLIDVAGAL 627

DB

QY 545 EQ--GAVYIFNGKFGGLSPQSRIQGAQVPPGIRWFGRSIHGKDLGDLADVVVGP 602

QY 628 GNAYILNSRPVQVQINASLHFPKSNITFHRCKSGSRDATCLAAFL--CFPTPIFLAPHF 685

DB

QY 603 GRVVLSRPVVDVVTLSFSPBEIPVHEVECSVSAREEQKHGKYLKACPKLPTPOQ 662

QY 686 -TTTVGIRYNATMDERRYPTRAHLDEGDRFTNRAVLSSQELCERINFHV-LDTADYV 743

DB

QY 663 GRLLANLSYTLQLDCHRMRSRGLPDPGSHLSGNTSI--TPDKSLDPHFPPICQDL 720

QY 744 KPVTFSEYSLEDPDHGPMDDG-----WPT--TLRVSVFPWNGCNEDEHCVPLDVL 794

DB

QY 721 SPINVLSNLSLEEGTPRDKQGRAMQFPLRPSIHTVTKEIPFERKNCDEKKCEANLTLS 780

QY 795 --ARSDLPAMEYQCVRLRPAQDCSAYTLSDFTVFIESTRQVAVAEATLENRGENAY 852

DB

QY 781 SPARSG-----PLRLMSGASL-----AVETLSNSGEDAY 810

QY 853 STVLNISOSANLQF--ASLIQ-----KEDSDGSIECTVNERRLOKQV-CHVSYPPF 900

DB

QY 811 WRLDLDFPRGLSPRKVEMLOPHSRMPVSCBELTEGS-----SLTKTKLKNVSSDIF 863

QY 901 RAKAVAPR-----LDSEFSKSIPLH--HLEIELAAGSDNSRDSKEDNVAFLRPHL 951

DB

QY 864 KAGQEVSLQVMFNTLLSSWEDFVELNGTVCE-----NENSSIQEDNSAATHPIV 914

QY 952 KYEADVL-----PT-----RSSLSH-YEVKLNSS-----LERYDGIGPPF 986

DB

QY 915 LYPNMLTKEQENSTLYISPTPKGKTQQVQVQVRIQPSAYDHNMTLEALYGVMPWH 974

QY 987 SCIFRIQNLGLFPIHGMWKITIPATRSNGLLKLRLDELDEANTSCNIWGNSTEVPT 1046

DB

QY 975 S-----EDPIT-----YTSVQTDPLVT 992

QY 1047 PVEEDLRAPOLNHSNDVWSINCNIRLVNQENFHLGLNML-RSLKALKYKSM--KI 1103

DB

QY 993 CHSBDLKR-PSSEAEQPCLPQVQPRCPVFRREILIQVTGTVLSKEIKASSTLSLCSL 1051

QY 1104 MVNAALQRFHSPFIPEEDPSROIEFEISKQEDQVPIWIVGSLTGLGLLALLVLLAL 1163

DB

QY 1052 SVSNSSKHPH---LYGSKASEAQLVKVDLIHSEKEM-LHVTYLSGGGLVLLFLFLAL 1107

QY 1164 RKLQFF-RSARRRR 1177

DB

QY 1108 YKVGFFKGNLKEKME 1122

RESULT 11

Q9R200

ID Q9R200 PRELIMINARY; PRT: 1160 AA.

AC Q9R200;

DT 01-MAY-2000 (TREMELrel. 13, Created)

DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)

DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)

DE Integrin alpha L.

GN ITGAL.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

QX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; Tissue=Spleen;  
RA Ma R.2., Teuscher C.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF065901; AAD25884.1; -  
DR HSP; P20701; ILPA.  
DR MGD; MGI:96606; Itgal.  
DR GO; GO:0008305; C:integrin complex; IEA.  
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.  
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP; 3  
DR Pfam; PF00357; integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1  
DR PRINTS; PRO1185; INTEGRINA.  
DR SMART; SM00453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS00234; VWA; 1  
SQ SEQUENCE 1160 AA; 128127 MW; A33C531B139FIFAD CRC64;  
  
Query Match 17.1%; Score 1063; DB 11; Length 1160;  
Best local Similarity 28.1%; Pred. No. 1.3e-64;  
Matches 358; Conservative 204; Mismatches 451; Indels 262; Gaps 55;  
  
QY 4 PRGLVAVAWL--SLWPGFTDTENMDTEKPRVPGSRTPAFPGYTVQOHDISGNKWLVVGAP 61  
DB 8 PRLLLLGLQFAKAW-----SYNLDTPQTSFLAQAGRHGYQLQIE-----DGVVVGAP 58  
  
QY 62 LETNGYKQTDGVVYKCPVIHGNCTKLNIGRVTLNVSERKDNMRGLSLATNPKNDFLAC 121  
DB 59 GEGD---NTGGLVHCRTSSPQQPVSLLH-----GSNHTSKYLGMTLATDAAGSILLAC 108  
  
QY 122 SPLMSHRCGSYYTTCMSRVNSNFRSKTV-----APALQRC-QTYMDIVIVLDGNS 173  
DB 109 DPLGSLRCDQNTYLSGLC-----YLPQSLGPMQNRPAQECMKGVLDVFLFDGSO 162  
  
QY 174 SI--YPWVEYOHPLINLKIFYIGPQOIQGVVQYGEDVVEHFLNDYRSVK-----DVV 226  
DB 163 SLQRKDFEKLFEKMDVWRK--LSNTSYQFAAQVQFDCRTEFTFLDY--VKQKQNPDLV 218  
  
QY 227 EAASHIQRGCTE-----TTPAGIEFARSEAF--QKGRKGAKKMIIVITDGESHSDPD 279  
DB 219 -----LGSVQPMFLITNTFRAINVYVAHVFKESGARPDATKVLVITDGERSDXGN 270  
  
QY 280 LEKVIQSERDNTVRYAVAVLGYNNRGINPDPETPLNEIKYIASDPDDKHFFNVTDK-AAL 338  
DB 271 I-----SAADHITRYIIGIGKHP--VSVQKQKTLH---IFASEPVEE-FVKILDTFEKL 318  
  
QY 339 KDIVDALGDRIFSEGTNKEB--TSFGLMSQTGFSSHVVEDGVLLGAVGYDNGAVLKE 397  
DB 319 KDLFTDLQRRIAYIEGTRNQLDTSFNMLSSSGISADLSKGHAVGVAGKDW----- 371  
  
QY 398 TSAGKVTPLRE-----SYLKEFPEELKNHGYLGYTTSVVVSQRQGVYVAGAPRNFHTG 452  
DB 372 --AGFLDLREDLQAGTFVQGEPLTSDVRGGLGYTGYVAMTSSRRPLLAAGAPRYQHVQ 429  
  
QY 453 KVILFTMN-----NRSITTHQAMRGQOIGSYFGSEITSDVIDDGDGVTDVLLVAGPYFN 507  
DB 430 QVLLFQAPAEAGRVNQNT---QKTEGTQIGSYFGGELCSVDLDQDGAELLIGAPLFFG 485  
  
QY 508 EGRERGKVVYVELPQNFYVNGTLKDSHTYQNAFSGSSIASVRDLNDSYNDVVVVGAPLE 567  
DB 486 EQR--GGRVFTYQRRQSLFEMVSELOQPGVPLGFGAAITALTDINDRLTDVAVGAPLE 544  
  
QY 568 DNHAGAIYIFHGFRGSLKTPKQRTASELATGQYFGCSIHGOLDNEDGLDLAVGAL 627  
DB 545 EQ--GAVYIFNGKPGGLSPQPSQRIQGAQVFPFGIRWFGRSIHGVKDLGGRLANVAVGPE 602

QY 628 GNAVILWSPVVOINASLHFEPSKINIPIHRCCKRGDRDATCLAAFL--CFTPIFLAPHFQ 685  
DB 603 GRVVVLSRPVVDVVTLSFSEPEEIPVHEVECSYSARBEQKHGVLKACFRIKPLTPQFQ 662  
  
QY 686 -TTTGVIRYNATMDERRYTTPRAHLEGGDRFTNRRAVLLSSGOELCERINHHV-LDTADTV 743  
DB 663 GRLLANLSYTLQLDGHRMERSGLPFGSHLSGNTSI--TPDKSCLDFHFPICIQDIL 720  
  
QY 744 KVPFSEVSELEDDPDHGMPLDDG-----WPT--TLRVSVPFMGNCNEDEHCVPLDVL 794  
DB 721 SPINVSLEFSLLEBEGTPDQGRAMQPIILRSIHVTKEIIFPKNCKEDKCEANLTL 780  
  
QY 795 --ARSDLPAMBYCORVLKPKQDCSAYTSLPDTTFFIESTRQRVAVATLENGENAY 852  
DB 781 SPARGG-----PLRLMSASL-----AVENTLNSGSDAY 810  
  
QY 853 STVLNLSQANLQF--ASLIQ-----KEDSGSIECVNEERRLOKQV-CNVSYPPF 900  
DB 811 WYRLDLPDPRGLSFRKVMQLPHSRPVSCBELTEGS-----SLLTYTLKCNVSSPIF 863  
  
QY 901 RAKAKVAFR-----LDSEFSKSIPLH--HLEIELAAGSDSNERDSTKEDNVAPLREHL 951  
DB 864 KAGQEVSLQVMENTLLNSWEDFVELNGTVHCE-----NENSSLOQEDNSAATHIPV 914  
  
QY 952 KYEADVL-----FT-----RSSLSH-YEVKLNSS-----LERYDGIQPPF 986  
DB 915 LIPVMLTKBQENSTLYISFTPKGPKTQOVQVHYQRIQPSAYDHNMPFLALVGVPRH 974  
  
QY 987 SCIFRIONLGLFPFHGMKMITIPIATRSNRLKLRLDPLTDEANTSCNINGSYRYET 1046  
DB 975 S-----EDLIT-----YTWVQTDPLVT 992  
  
QY 1047 PVEDILRAPQLNHSNDSVVSINCLIRLVNQNFINPHLLGNLWL-RSLKALKYKSM--KI 1103  
DB 993 CHSEDLKRPSSBAEPLCGVQPRCPF--VFRWEILIQVTGVELSKBIKASSTLSCSSL 1050  
  
QY 1104 MYNAALQROFHSPFIFREEDPSRQIEFEISKOEDQVPIIIVGSTGLGLLLALLALVAL 1163  
DB 1051 SVSENSSKHFH--LYGSKASEAQVLVKVLDLHKEM-LEVVVLSGIGLVLLFLIFLAL 1106  
  
QY 1164 RKLGF-RRARRR 1177  
DB 1107 YKVGFFKRYLKEKE 1121  
  
RESULT 12  
Q98TF0 PRELIMINARY; PRT; 1187 AA.  
AC Q98TF0;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE CL11-2.  
GN Cyprinus carpio (Common carp).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Cyprinus.  
OX NCBI\_TaxID=7962;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Peritoneal exudate cells;  
RA Kimura M., Fujiki K., Nakao M.;  
RT "Molecular cloning of a leukocyte integrin from the common carp."  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB048537; BAB39135.1; -  
DR HSP; P20701; ILPA.  
DR GO; GO:0008305; C:integrin complex; IEA.  
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.  
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001963; Asprotease\_AS.



```
Db 1 VQAYDWSGTVVHESKQSFPP-----YKAFKVLHDBNQSSVYGVSV-AVINKNKSVHP 54
QY 442 VAGAPRNTKGVILFTMHNRSLTHQAMRGQIGSYFGSEITSDVDIGDGVTVLLVG 501
Db 55 VAGAPRTDYGQVWVYVITSGQVNSVITQORGEQIGSYFGSVLCSVDNRDSITDVLVG 114
QY 502 APWYFNE-GRERGKVVYVLRQNRFPVYVNGTLKDSHSYQNAFGSSIASVRDLNODSYNDV 560
Db 115 APTFMNEYKEEGOVYMFSDGILVQREQLGPKSLENTRFQSAIVELSDIDLQYNDV 174
QY 561 VQAPLEBNNHAGAIYIPHGFRGILKTPKRIITASELATGLQVFGGSIHQGLDLNBDGLI 620
Db 175 IVGAPLEQNSGAIYIPNGKNTIRTKYSKILGSSFPNGLOVFGRSVDGHQDLNGDTIN 234
QY 621 DLAVGALGNVILWSPVVOINASLHPEPSKINIFHRDCKRSORDATCLAAFCPTPIEL 680
Db 235 DVSVGAYGVQILWSSHVADVSVKALFIPQOIVLTNNKAEVTVR-----ICFSAVP- 285
QY 681 APHFOTTTGIRVATND-----RRYTPRAHLEGGDRFTNRAVLSSQOELCERINFV 736
Db 286 RPANSQNSVAVTNTATLDALLSRVTSRQGFRESNDRFLQRNIVIGCTPESCFEHV-FNV 344
QY 737 LOTADYVXPTVPSEVSELEDDHGMVLDGMPITLVSVPFWNGCKNEDEHCVPLVLDAR 796
Db 345 QETSDSENALRINVAQPGSSPVLNYPSSSTENFPLKDCGDKICISDLSLQIQ 404
QY 797 SDLPAMEYQORVLRPAQDCSAYTLSPDTVTPIIESTQRVAEATLENRGENAYSTVL 856
Db 405 -QIPSDIK-----SPVIVSKDRLLFQVTLNKLNAENYEL 441
QY 857 NISQSANLOPAS-----LLOKEDSGSTECVNEERLOKQVNVSPFPFRKAKV 906
Db 442 NATPSENLFASSTIPRDTGTEVLCQVNTQNSVCL-----IGFPELKDQSKV 489
QY 907 APRLDSFYSKIFLHLEIBLAAGSDSNRDSYKEDNVAPLRPHLYEADVLFTRSSLS 966
Db 490 IPDIWFENLNNQNTAFIVFQATSES--REAREEDNSVNTLPVQYDAEIHILTRVTIN 547
QY 967 HYEV-----KLNSSLEYDGIQPPSCIFRIONGLPFIHGMMKIPIATSGNRLKL 1022
Db 548 FYEVFPLGNIPSVVNTFDDIGLEFNFTVKV-STGYPLKTAVLSINVPQFTKSNPLLYI 606
QY 1023 RDLPTDE-ANTSCNINWGNSTE-----YRPTPVEEDLRAPQLNHSNSDVVSYNC---NIR 1073
Db 607 TSVQTDQMTGVSSNRVINPAHVEKIPYKPSFKEENLKAVHVNCSARCAVKCFLENLQ 666
QY 1074 LVPNQEFHLLGNLWLSLKALKYSKIMVNAALQROFHSFPIFREDDPDRQIEFFIS 1133
Db 667 LREBYTVN--VSTRINWGTFAAAHFTQIELVANAKLETIDSEIFVVSNDV--MPLTIT 722
QY 1134 KQ-EDQVPTIIVGSTLGLLALLLVLALRLKGFPR 1170
Db 723 KPGKSEVPIGVVIGSAIAGLVALLVALLWKLGFKK 760
```

## RESULT 14

```
Q96HB1 PRELIMINARY; PRT; 1086 AA.
ID Q96HB1;
AC Q96HB1;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; BC005777; AAH08777.1; -.
DR GO; GO:0008305; C:inhibitor complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00992; vwa; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWF A; 1.
DR Hypothetical protein.
KW SEQUENCE 1086 AA; 119223 MW; F6PF2546E8C632F9 CRC64;
```

Query Match 16.4%; Score 1018.5; DB 4; Length 1086;  
Best Local Similarity 27.8%; Pred. No. 1.4e-61;  
Matches 350; Conservative 184; Mismatches 420; Indels 303; Gaps 53;

```
QY 8 VVAMALSLWPGP-----TDTNMDTRKPRVIPGSRTA-PFGYTVQOHDISGNKMLVVGAP 61
Db 6 ITVAMALLSGPFPFAPASSYNLDVGRGSRSPPRAGRHFGYRVLQ---VGH-GVIVGAP 61
QY 62 LETNGYQKTDVYKCPVIHGMCTKLNLRVTL--SNVSEKDNMRLGLSLATNPKNQNSFL 119
Db 62 GEGN---STGSLYQCQSGTGHCLP-----VTLRGSNYTSK---YLGMTLATDPTDGSIL 109
QY 120 ACSPLWSHEGSSYYTTCMCSEVSNFRFSKTVAPALQRCQTYMDIVVLDSGNSIYPMV 179
Db 110 FAA----- 112
QY 180 EYQHFELNLIKFFYIGPQIQVGVVQYGEDVVEHFLNDYRSVKDWEAAASHIEQGGTE 239
Db 113 -----VQSTSYKTEPFDSDYVYKRPDALLKHVXEM-LLL 147
QY 240 TETAPGIEFARSEAFQK--GGRKGAKKYVITGESHDSPLDLEKVIQSSRDNYTVAY 297
Db 148 TMTFQAINVATEVREELGARPDATKVLIIITDEATDSGNIDAA-----KDLIRII 201
QY 298 AVLGYNRRGINPETFLNEIKYIASDPDRKFFNVTD-AALKDVIDALGRDIFSLGNT 356
Db 202 GI-GKHFTKESQET---LHKFASKPASE-FVKILDTFEKLKDLFTELQKIVYIEGTS 255
QY 357 KYE-TSFGLEMSQTGFSHHVEDGVLLGAVGAYDNGAVLKETSGKVIPLR-----ESY 410
Db 256 KODLTSFMELSSSGISADLSRGHAYVGVAGAKDM-----AGGFLDLKADLQDDTF 306
QY 411 LKEPPEELKNHGAIVGYTVTVTSVSRQGRVYVAGAPRNFHTGKVLFTM-----HNNRSL 465
Db 307 IGNEPLTPEVRAGLYGYTVTLPSRQKTSLLASGAPRYOHMGRVLLFQEPQGGHWSQV 366
QY 466 TIHQMRQQQIGSPFGSEITSDVDIGDGVTVLIVGAPMYFNEGRERKQVTVYELQRNP 525
Db 367 TIH-----GTQIGSYFGGELCGVDVQDGETELLIGAPLFYGEOR-GGRVFIYORRQIGF 421
QY 526 VYNGTLKDSHSYQNAFGSSIASVRDLNODSYNDVWVWVWVWVWVWVWVWVWVWVWVWVW 585
Db 422 EEVSELOQDPGYPIGRFGEAITALTDINGDLVDVAVGAPLEEQ--GAVYIFNGEKGCLS 479
QY 586 KTKPQRITASELATGLQVFGGSIHQGLDLNBDGLIDLAVGALGNVILWSPVVOINASL 645
Db 480 PQPSQRIEQTQVLSGIQWFGRSIHGVKXDEGLDGLADVAGAESQMVILSSRPVVDVMTLM 539
QY 646 HFESKINIPIHEDCK-----RSGRDATCLAAFLCTPIPLAFHFOFTTV-GIRYNATM 697
Db 540 SPSFAETPVHEVCSYSTSNOKKEGVNIT-----ICFOIKSLIIPQOQLVANITYTLQL 594
QY 698 DERYTTPRAHLEGGDRFTNRAVLSSQOELCERINFHV-LDTADYVXKVPVTFSEVSELD 756
Db 595 DQHR-TRERGLFPGGRHELRENIAVTTSMSC-CTDFSHPFVQVQDLISPINVSLNFSLWE 652
```

```
QY 757 PDHGP---MLDGMPTFLAVS-----VPEWNGCNEDEHCVPLVLDARSDLPTAMEVC 806
D 653 EGTREDQAGKDIPILRPSLHSETWEIPFENKCEKKEANL----- 697
QY 807 QVILKAPQDCSAYTLSPOTTVIIESTQRVAVATLENRGENAYSTVLTNSQSANTQF 866
D 698 -RVSPSPARSARLRLTAFAS-----LSVELSLNLEEDAYVWQLDLHFPPLGSP 745
QY 867 ASL-TOKEDSDGSIEC--VNEERRLQKOV--CNVSPFPRKAKVAFRL-----DSBFS 915
D 746 RVEMLKPHSQIPVSCSEELPEBSRLLSRALSNCVSPPIFKAGSHVALQMMFTLVNSWG 805
QY 916 KSIFLHLEIEIAGAOSNERSTKEDNVAPLRFHLKYEADVL-----FTRSS 963
D 806 DSVELH-----ANVTCCNEDSLLDNEASATTIIPILYPINILIOQEDSTLYVFTPKG 859
QY 964 SLUSH-----YEVKLANS-----LERYDGIQPPFSCIFRQNLGLPIHGMKMIIP 1010
D 860 PKIHQVKKMYQVRIQPSIHDHNIPTLEAVVGPQP-----PSEG-----P 899
1011 IATRSNRLLLKRLDFTLDBANTSCNIWGNSTGYRTPVE-EDLRAPQLNHSNDVWSIN 1069
D 900 ITHQ-----WSVQMB-PPVPCHYEDLERLPDAAEPCFALFR 936
QY 1070 CNIRLVNQEINFHLLGNLW-----RLKALKYKSMKIMNALQRFHSPPIFRBED 1123
D 937 CPV--VFROEILVQVIGTLELVEIEASSMFL-CSSLISISFNSS--KHFF---LYGSNA 988
QY 1124 PSRQIEFEISKEDQVPTIWIIVGSLGILLALLLALRLXIGFF-RSARRRRPPG 1179
D 989 SLAQVVMKVDVYVEXQM-LYLVLSGIGSLLILLILFVLYKGFKNLKEKMEAG 1044

RESULT 15
O88341 PRELIMINARY; PRT; 1167 AA.
ID O88341
AC O88341
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DE Integrin alpha E2 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98056820; PubMed=9394838;
RA Brenan M., Rees D.J.G.;
RT "Sequence analysis of rat integrin alpha E1 and alpha E2 subunits:
RT tissue expression reveals phenotypic similarities between
RT intraepithelial lymphocytes and dendritic cells in lymph.";
RL Eur. J. Immunol. 27:3070-3079(1997).
DR EMBL; AF020046; AAC23663.1; -.
DR HSSP; P11215; 1BHQ.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PSS0234; VWEA; 1.
FT NON TER.
SQ SEQUENCE 1167 AA; 128593 MW; E3FED7E18B5CBBEC CRC64;
```

```
Query Match 15.7%; Score 979; DB 11; Length 1167;
Best Local Similarity 26.3%; Pred. No. 9e-59;
Matches 337; Conservative 212; Mismatches 479; Indels 252; Gaps 48;

QY 10 AWALSILPFGTDTTNDTRKPRVPGSKTAFGGTVVQOHDISGKNKLVVGPAPLETNGYQK 69
D 25 SWVTALQPGAPSVLN-----SLHQDFSNQTCILLVARRSN--RN 63
QY 70 TGVVYKCVIHGNCCKL--NLGRVTLNVSERKDNMLGLSLATNPKNSFLACSPWS 126
D 64 SAVIYQC-----ATSIIPDIBGQPVHILMPKRVQ-GVTLVRN--HNGVLVCIQVS 114
QY 127 HECGS-SYTTTGMCSRNSNF----- 146
D 115 RKPSRLNSELTCACSLTTPNLDLQAYFSDLEGVLGLGASVNSGDYVSGSGSTGEETK 174
QY 147 ---RSKTVAPALQCQTYMDIVIVLGSNSIYP--WVEVQHFLINILKVIYIGQIQV 201
D 175 SARLRQAVBEEBEEAGTBIIVLDGSGSIESDFQAKDFISTWNRNYEKFCECNF 234
QY 202 GWVOYGEDVVEHFFHLNDYRSVKVVEAASHIEQRGGTETRTAFGIEFARSAP--QKGR 259
D 235 ALVQYGGVIQTEFDLLDSRDINASLAKVQSIQVQKEV-TKTASAMQHVLNDNIFPSRGR 293
QY 260 KGAKYMWIVITDGESHSP-DLEKVIQOSERDNTTRAVAVLGYNRRGINPETFLNIX 318
D 294 KALKVMVVLTDGDFIDPLNLATTVISSKQGVVRAIGVGNAPF-----NNNTY-RELK 348
QY 319 YIASPDQDKHFNVTDEAALXDIVDALCDRIPLSELTGNKNETSGLEMSQTFGSHVVED 378
D 349 LIASDPKAAHTFKVTNYNSALDGLLSLQQRILHMEGTVDLQY--QLAQTFSAQILDK 406
QY 379 G-VLLGAVGAYDMN-GAVLKETSAGKVIPLRESYLKEFPPELKNHGYLGTVTTSVSSR 436
D 407 CQVLLGTGAFNWSGGALLYNTQNGRGRFLNQAKEDFRAA---QYSYLGYSV-AALHKA 462
QY 437 QGRVYVAGAPRNFHTGKIVILFTMNNRSLTTHQAMRGQOIGSYFGSETTSVIDGDGVD 496
D 463 HGVSIVAGAPRKLKRGAVFELQKEDGEE-TPMRRIESEQMSYFGSVLCPVDINWDGTD 521
QY 497 VLLVGAPMYFNEGREGKVVYVELRQN----RFVYNGTLKDSHSHYQNAFSGSIASVRL 552
D 522 PLLVAAPFYHIRG-EEGRVTVYRVHEQDAPPSLVY--TLSGYPGLTSSRFQFAMAAGDI 578
QY 553 NODSYNDVNVGAPLEDNEA-----GAIYFPHGFRGSLTKTPKORITASELATGQYFGC 606
D 579 NQDKFTDVAIGAPLEGFGAGDGSYGVYINGHSGGLHASPSQOIRASSVALGLYFPM 638
QY 607 SIHQDLNEDGLIDLAVGALGNVILWSRPVQINASHLHFEPSKINIPIHRDCKRSGDA 666
D 639 SVSGGLDFSGDDLADITVGSQDVAVLRSRPVVDLTVSMFTPDALPMAFKD----- 690
QY 667 TCLAAFLCFPIFLAPHFQTTTCIRYNATWDERRYTPRAHLD-----EGG 712
D 691 -KMDVBLCFKVDSSAVPSEPLRGMSLNFVTDVDTQKORLOQCADRSCCSCMLKWSGG 749
QY 713 DRFTNRVAVLASSGOBLERINFHVLDTADYVVKVPTFSVYSLEDPDH-----GPMLDGW 767
D 750 SSLCEHFCGLISTEEELCS-----DDCFSNITIKVSYEFQTSERRNHPNFIIDHYK 800
QY 768 PTLURVSVPTWNGCNEDEHCVPLDVLDAESDLPTAMEYCORVLKRPQDCSAYTLSDFT 827
D 801 EPSAIFQLPYEKDCKNVCIAEI-----QLTTAIS-----QOD----- 834
QY 828 VFIIESTRQVAVEATLENGENAYSTVLTNSQSANLQFASLQKSDSGSIECVNBERR 887
D 835 --LVVGITKEVTMMNISLTNSGDSYMTNMAINPRNLQFKK-IQKPLSP-DIQCDPFPV 890
QY 888 LQKOV--CNVSPFPRKAKVAFELDSFESKSTFLHLEIEIAGAOSNERDSTKENVA 945
D 891 ASVLWACKIGHPIK-RSSVNVSVTWQLBESIPNRTADITVTISNKSLEARETH-- 947
QY 946 PLRFLHLYEA-----DVLFP--TRSSLSLHYEVKLNSSLERYDVGIGPPSPCFRIQLGLF 998
```



Db 948 SLOFRHAFI AVL SRPSVTWNTSQSSSDHKEFFNVHGENH-----FGAVFOLQ----- 996  
QY 999 PIHGMMKTIPIATRSNRLKLRDPLTDEANTSCNNGNSTEYRPTFVEEDLRAPOL 1058  
Db 997 -----ICVPITLRD-LQIRVKHLTKTQAHTEC-----TQSQBPTCGSDPVQNVVEW 1042  
QY 1059 NH-----SNSDVVSINCNIRLVPNOEINFHLLGNLWLRSLKALKY-----KSMKIMV 1105  
Db 1043 HSVICAITSNKENVTVAEISMGHTKQL-----LRDISLQILGEISFNKSLYEG 1093  
QY 1106 NAALQRFHSPFIIFREEDPSRQIEFEISKQEDWQVFIMIVGSLTGLGLLLALLVLALRK 1165  
Db 1094 NAENHRTKITVIFLKEEKPH-----SLP-LIIGSSIGLLVLVVIILFK 1138  
QY 1166 LGHFRSARRRREPGLDPTPK 1185  
Db 1139 CGPPK--RKYQLNLESVRK 1156

Search completed: June 24, 2004, 17:55:41  
Job time : 64 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 17:53:06 ; Search time 64 Seconds  
(without alignments)  
5244.788 Million cell updates/sec

Title: US-09-980-403-2  
Perfect score: 6224  
Sequence: 1 MDLPRGLVVALSIWPGFT.....FRSARRRRRPGDTPKVLK 1188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : A.Genesep\_29Jan04:\*
- 1: Genesep1980s:\*
  - 2: Genesep1990s:\*
  - 3: Genesep2000s:\*
  - 4: Genesep2001s:\*
  - 5: Genesep2002s:\*
  - 6: Genesep2003as:\*
  - 7: Genesep2003bs:\*
  - 8: Genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6224	100.0	1188	4	AAB30929 Amino aci
2	6196	99.6	1188	4	Aau14231 Human nov
3	6196	99.6	1188	4	Aab50085 Human A25
4	6196	99.6	1188	5	Aau10551 Human A25
5	6192.5	99.5	1189	4	Abg12949 Novel hum
6	6192	99.5	1188	4	Aau14467 Human nov
7	6192	99.5	1188	7	Ade09956 Novel pro
8	6188.5	99.4	1189	3	Aab25582 ITGAl1 pr
9	6188.5	99.4	1189	6	Abr58364 Human NOV
10	6188.5	99.4	1189	6	Ada27054 Human nov
11	6188.5	99.4	1189	7	Ade63570 Human Pro
12	6188.5	99.4	1189	8	Ade8584 Novel hum
13	5780	92.9	1120	6	Abr58365 Human NOV
14	5693	91.5	1188	4	Aab50087 Murine A2
15	5693	91.5	1188	5	Aau10552 Murine A2
16	5383	86.5	1034	3	Aab25590 Protein e
17	5383	86.5	1034	6	Ada27062 Human nov
18	5383	86.5	1034	8	Ade8592 Novel hum
19	3621	58.2	707	4	Aau19663 Human nov
20	3621	58.2	707	5	Abp47883 Human pol
21	3621	58.2	707	7	Adc10845 Human ext
22	3182	51.1	688	5	Abb72300 Rat prote
23	3113	50.0	696	5	Abb72289 Rat prote
24	2518	40.5	545	5	Abb72288 Murine pr
25	2403.5	38.6	1167	3	Aay32242 Human int

26	2401.5	38.6	1152	4	AAB64657	Aab64657 Human sec
27	2375.5	38.2	1167	4	AAB64584	Aab64584 Human sec
28	2375.5	38.2	1167	6	ABP99490	Abp99490 Human gen
29	2375.5	38.2	1167	6	ABR00964	Abr00964 Human gen
30	2375.5	38.2	1167	6	ADA44026	Ada44026 Human sec
31	2373.5	38.1	1152	4	AAB64658	Aab64658 Human sec
32	2319	37.3	1132	3	AAY32243	Aay32243 Human int
33	2166.5	34.8	1179	5	ABB90759	Abb90759 Human Tum
34	2166.5	34.8	1179	5	ABP64915	Abp64915 Human pro
35	2166.5	34.8	1179	5	ABUS4466	Abu54466 Human int
36	2159	34.7	1151	8	ADE86652	Ade86652 Human int
37	2149.5	34.5	1180	5	ABB90788	Abb90788 Rat Tumou
38	2149.5	34.5	1180	6	ABU54495	Abu54495 Mouse tum
39	2149.5	34.5	1180	7	ADE63568	Ade63568 Rat Prote
40	2093	33.6	1183	2	AAy07728	Aay07728 Armenian
41	1863	29.9	1181	6	ABU03548	Abu03548 Angiogene
42	1863	29.9	1181	6	ABRS9703	Abr59703 Human via
43	1863	29.9	1181	6	AAG97975	Aag97975 Alpha2 in
44	1863	29.9	1181	6	ABU03616	Abu03616 Human exp
45	1863	29.9	1181	6	ABU03614	Abu03614 Human exp

ALIGNMENTS

RESULT 1  
AAB30929  
ID AAB30929 standard; protein; 1188 AA.

XX AAB30929;

AC AAB30929;

DT 02-APR-2001 (first entry)

XX Amino acid sequence of a human alpha11 integrin chain.

KW Human; integrin; alpha11 subunit; fibroblast; muscle cell; chondrocyte;

KW osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;

KW wound healing; trauma; rheumatoid arthritis; osteoarthritis;

KW osteoporosis; cartilage damage; bone damage; cartilage.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..22

FT Region /note= "signal peptide"

FT Domain /note= "leucine zipper"

FT Domain 1142..1164

FT Domain /note= "transmembrane domain"

XX WO2000075187-A1.

XX 14-DEC-2000.

XX 31-MAY-2000; 2000WO-SR001135.

XX 03-JUN-1999; 99SE-00002056.

XX (ACTI-) ACTIVE BIOTECH AB.

XX Gullberg D;

XX WPI; 2001-071061/08.

XX N-PSDB; AAC86871.

XX Integrin subunit alpha 11 or integrin heterodimer comprising subunit

PT alpha 11 in association with subunit beta, useful for treating muscle

PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.

XX Disclosure; Fig 2a-c; 79pp; English.

XX The present sequence represents a human integrin subunit, designated

CC alpha11. The alpha11 polynucleotide and polypeptide are useful as markers





FT Domain 478..516  
 FT /label= Integrin\_alpha subunit\_repeat\_domain\_#5  
 FT Domain 540..575  
 FT /label= Integrin\_alpha subunit\_repeat\_domain\_#6  
 FT Domain 602..640  
 FT /label= Integrin\_alpha subunit\_repeat\_domain\_#7  
 FT Domain 1142..1164  
 FT /label= Transmembrane\_domain  
 FT Domain 1165..1188  
 FT /label= Cytoplasmic\_domain  
 XX W0200073339-AL.  
 XX 07-DEC-2000.  
 XX 15-MAY-2000; 2000WO-US013262.  
 XX 28-MAY-1999; 99US-00322790.  
 PR 27-APR-2000; 2000US-00561263.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA Pan Y, Lora JM;  
 PI WPI: 2001-041142/05,  
 DR N-PSDB; AAC91901, AAC91902.  
 XX Nucleic acid encoding alpha-integrin subunits, useful for treatment and  
 PT diagnosis of fibrosis, e.g. of the liver.  
 XX Claim 8; Fig 1; 164pp; English.  
 CC The present sequence is human integrin alpha subunit, A259, A259 is  
 CC homologous with the alpha1 and alpha10 integrin subunits and is  
 CC overexpressed in fibrosis. A259 is implicated in regulation of  
 CC proliferation, differentiation and/or function of many different cell  
 CC types. Inhibitors of A259 activity are useful for the treatment of liver  
 CC disease, particularly fibrosis, and also fibrosis in other organs  
 CC (specifically lung and kidney). In addition, A259 can be used for  
 CC treatment and prevention of cancer, osteoporosis, acute myeloid  
 CC leukaemia, HIV infection, and rheumatoid arthritis  
 XX Sequence 1188 AA;  
 SQ  
 Query Match 99.6%; Score 6196; DB 4; Length 1188;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 1184; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MDLPRLGVAVALS LWPFGFTDTFMDTRKPRVPGSRTAPFGYTVQOHDISGNKMLVWGA 60  
 DB 1 MDLPRLGVAVALS LWPFGFTDTFMDTRKPRVPGSRTAPFGYTVQOHDISGNKMLVWGA 60  
 QY 61 PLETNQYKTDYVYKCPVHNGCTKLNLRVTLNSYSEKDNKRLGLSLATNPXNSFLA 120  
 DB 61 PLETNQYKTDYVYKCPVHNGCTKLNLRVTLNSYSEKDNKRLGLSLATNPXNSFLA 120  
 QY 121 CSPLWHECGSSYTTGMSRVNSNFRFSKTVAPALQRCQTYMDIIVLDGNSNIPWVE 180  
 DB 121 CSPLWHECGSSYTTGMSRVNSNFRFSKTVAPALQRCQTYMDIIVLDGNSNIPWVE 180  
 QY 181 VQHFLNLLKFKYIGQIQGVVQYGEDVHFEHFLNDYRSVNDVVEAASHIQRGGTET 240  
 DB 181 VQHFLNLLKFKYIGQIQGVVQYGEDVHFEHFLNDYRSVNDVVEAASHIQRGGTET 240  
 QY 241 RTAFGIEFARSEAFQGGKRGAKVMIVITDGSHPDPLEKVIQOSERDNTVRYAVVL 300  
 DB 241 RTAFGIEFARSEAFQGGKRGAKVMIVITDGSHPDPLEKVIQOSERDNTVRYAVVL 300  
 QY 301 GYNRGINPETTLEIKYIASPDQKHPNVTDEALXDIIVDALGDRIFESLEGTNKNET 360  
 DB 301 GYNRGINPETTLEIKYIASPDQKHPNVTDEALXDIIVDALGDRIFESLEGTNKNET 360  
 QY 361 SFGLEMSQTGFSSHHVVDGVLGAVGAYDWMNGAVLKETSAGKVIPLRESYLKEFPEELKN 420

DB 361 SFGLEMSQTGFSSHHVVDGVLGAVGAYDWMNGAVLKETSAGKVIPLRESYLKEFPEELKN 420  
 QY 421 HGAYLGYTVTSVVSRRQGRVYVAGAPRNNYTKVILFTWNNRSLTIHQMGQOIGSYF 480  
 DB 421 HGAYLGYTVTSVVSRRQGRVYVAGAPRNNYTKVILFTWNNRSLTIHQMGQOIGSYF 480  
 QY 481 GSEITSVDIDGVDVLLVAGAPMYFNEGREGKGVYVYELRQNRVYVYAGTLDKDSHYQNA 540  
 DB 481 GSEITSVDIDGVDVLLVAGAPMYFNEGREGKGVYVYELRQNRVYVYAGTLDKDSHYQNA 540  
 QY 541 RGSSTIASVRDLNQDSYNDVVGAPLEONHAGAIYIFHGFRGSIILKTPKORITASELATG 600  
 DB 541 RGSSTIASVRDLNQDSYNDVVGAPLEONHAGAIYIFHGFRGSIILKTPKORITASELATG 600  
 QY 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILNSRPVQINASLHFPSPKINIFHRDCK 660  
 DB 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILNSRPVQINASLHFPSPKINIFHRDCK 660  
 QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYXATMDERYTPRAHLDGGRDTRNAV 720  
 DB 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYXATMDERYTPRAHLDGGRDTRNAV 720  
 QY 721 LLSQQLCERINPHVLTADYVKPVTSEVYSLEDDPDHGMKLDGMPPTLRVSVPFWNG 780  
 DB 721 LLSQQLCERINPHVLTADYVKPVTSEVYSLEDDPDHGMKLDGMPPTLRVSVPFWNG 780  
 QY 781 CNEDEHCVPDLVDARSDLPTAMEYCORVLRKPAQDCSAYTILSPOTTFVILFESTQRVAV 840  
 DB 781 CNEDEHCVPDLVDARSDLPTAMEYCORVLRKPAQDCSAYTILSPOTTFVILFESTQRVAV 840  
 QY 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGSIQECVNEERRLOKQVCNVSYPFF 900  
 DB 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGSIQECVNEERRLOKQVCNVSYPFF 900  
 QY 901 RAKAVAPRLDSEFSKSIPLHLHLEIELAAGSDNSRSTKEDNVAPLPHLKYEADVLPT 960  
 DB 901 RAKAVAPRLDSEFSKSIPLHLHLEIELAAGSDNSRSTKEDNVAPLPHLKYEADVLPT 960  
 QY 961 RSSLSHVEVKLNSLERYDGIQPPSCIPRIQNLGFLPIHGMKMIPIPIATRSNRL 1020  
 DB 961 RSSLSHVEVKLNSLERYDGIQPPSCIPRIQNLGFLPIHGMKMIPIPIATRSNRL 1020  
 QY 1021 KLRFPLTDEANTSCNMGNSNTEYRPTVEEDLRAPQLNHSNSDVVSINCMRLVNPQEI 1080  
 DB 1021 KLRFPLTDEANTSCNMGNSNTEYRPTVEEDLRAPQLNHSNSDVVSINCMRLVNPQEI 1080  
 QY 1081 NFHLGNLWLSLKLKLYSKMKIMVNAALQRFSPFIREDPSRQIEFEISKOEDQV 1140  
 DB 1081 NFHLGNLWLSLKLKLYSKMKIMVNAALQRFSPFIREDPSRQIEFEISKOEDQV 1140  
 QY 1141 PIWIIIVGSTLGGLLALLLALLLALRLKGLFPRSARRRREPGLDPTPKVLE 1188  
 DB 1141 PIWIIIVGSTLGGLLALLLALLLALRLKGLFPRSARRRREPGLDPTPKVLE 1188  
 RESULT 4  
 AAU10551  
 ID AAU10551 standard; protein; 1188 AA.  
 XX  
 AC AAU10551;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Human A259 polypeptide.  
 XX  
 KW Human; A259; integrin alpha subunit; integrin alpha 10; secreted protein;  
 KW liver disease; fibrosis; lung; kidney; bone associated disorder; blood;  
 KW cartilage associated disorder; hematopoietic disorder; bone marrow;  
 KW immune related disease; apoptotic disorder; neuronal tissue disease;  
 KW neurodegenerative disease; gene therapy; cancer; cytostatic; osteopathic;  
 KW nephrotropic; immunomodulator; anti-inflammatory; neuroprotective;  
 KW antiarthritic; antianaemic; antiallergic; antiasthmatic; dermatological;

antidiabetic; anticonvulsant; antiparkinsonian.

Homo sapiens.

Key Location/Qualifiers

1..1141 /note= "Extracellular domain"

1..22 /note= "Signal peptide"

23..1188 /note= "Mature human A259"

37..90 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"

115..157 /note= "Integrin alpha repeat domain"

164..345 /note= "I domain or Von Willebrand Factor type A domain"

367..392 /note= "Integrin alpha repeat domain"

421..472 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"

476..532 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"

538..593 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"

600..654 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"

1142..1164 /note= "Transmembrane domain"

1165..1188 /note= "Cytoplasmic domain"

WC200181414-A2.

01-NOV-2001.

27-APR-2001; 2001WO-US013516.

27-APR-2000; 2000US-00561263.

(MILL-) MILLENNIUM PHARM INC.

Pan Y, Lora J;

WPI; 2002-041397/05.

N-PSDB; AAS16873.

New A259 nucleic acids and polypeptides, which comprise integrin alpha subunit, useful for diagnosing, preventing or treating e.g. liver disease, kidney or lung fibrosis, cancers, blood disorders or immune related diseases.

Claim 9; Fig 1; 168pp; English.

The invention relates to human and murine A259 nucleic acid molecules which encode secreted proteins with homology to integrin alpha subunits. Specifically to integrin alpha 10. The A259 polypeptide and nucleic acid are useful for treating liver disease or fibrosis, particularly kidney fibrosis or lung fibrosis. The A259 polypeptide and nucleic acid are also useful for diagnosing, preventing or treating cartilage and bone associated disorders (such as bone cancer, achondroplasia, myeloma, fibrous dysplasia, scoliosis, osteoarthritis, osteosarcoma and osteoporosis), bone marrow, blood and haematopoietic disorders (such as acute myeloid leukaemia, haemophilia, anaemia and thalassemia), immune related diseases (such as HIV, viral infections, cancers, T cell autoimmune disorders e.g. AIDS, and allergic inflammatory disorders e.g. asthma and psoriasis), apoptotic disorders (such as systemic lupus erythematosus and insulin-dependent diabetes mellitus), diseases of the

CC neuronal tissues (such as epilepsy and muscular dystrophy) and  
CC neurodegenerative diseases (such as Parkinson's disease and Huntington's  
CC disease). This sequence represents the human A259 polypeptide  
XX  
SQ Sequence 1188 AA;

Query Match 99.6%; Score 6196; DB 5; Length 1188;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1184; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDLPRGLVAVKALSLMPGFTDTFNMDTRKPRVIPSRTAFPGYTVQOQHDISGNKWLTVGA 60  
DB 1 MDLPRGLVAVKALSLMPGFTDTFNMDTRKPRVIPSRTAFPGYTVQOQHDISGNKWLTVGA 60  
QY 61 PLENGYQKTDGVYKCPVHGNCTKLANGRVTLNVSERKDNMRIGLSLATNPKDNSPLA 120  
DB 61 PLENGYQKTDGVYKCPVHGNCTKLANGRVTLNVSERKDNMRIGLSLATNPKDNSPLA 120  
QY 121 CSPLWSHECGSSYYTTCMCSRVSNSFRPSKTVAPALQCTQYMDIVIVLDCSNSIYPWVE 180  
DB 121 CSPLWSHECGSSYYTTCMCSRVSNSFRPSKTVAPALQCTQYMDIVIVLDCSNSIYPWVE 180  
QY 181 VQFLINILKXFFYIGPGQIQGVVQYGEDVVHFEHLNDYRSKDVVEAASHIEQGGTET 240  
DB 181 VQFLINILKXFFYIGPGQIQGVVQYGEDVVHFEHLNDYRSKDVVEAASHIEQGGTET 240  
QY 241 RTAGIEFARSEATQKGRKCAKVMIVITDGRSHSDPLEKVIQOOSERDNRVAVAVL 300  
DB 241 RTAGIEFARSEATQKGRKCAKVMIVITDGRSHSDPLEKVIQOOSERDNRVAVAVL 300  
QY 301 GYNNRRGINPETFLEINIKYIASDDPDKHFNFVNTDEAALKOIVDALGDRIFSLGNTNKNET 360  
DB 301 GYNNRRGINPETFLEINIKYIASDDPDKHFNFVNTDEAALKOIVDALGDRIFSLGNTNKNET 360  
QY 361 SFGLEMSQTFSSHVVDGVLIGAVYDNGAVLKETSAGKVIPLRESYLKPFEEELKN 420  
DB 361 SFGLEMSQTFSSHVVDGVLIGAVYDNGAVLKETSAGKVIPLRESYLKPFEEELKN 420  
QY 421 HGAVLGTVTSVSSRQGRVYVAGAPRNFHTGKVIPLFMHNRSLTIHQMRGQOIGSYF 480  
DB 421 HGAVLGTVTSVSSRQGRVYVAGAPRNFHTGKVIPLFMHNRSLTIHQMRGQOIGSYF 480  
QY 481 GSEITSVDIDGVDVLLVGA PMYFNEGRGRGVYVYELQNRFPVYNGTKDKSHSYQNA 540  
DB 481 GSEITSVDIDGVDVLLVGA PMYFNEGRGRGVYVYELQNRFPVYNGTKDKSHSYQNA 540  
QY 541 RFGSSIASVRDLNODSYNDVVGAPLENDHAGALVIPHGRGSLTKTPKQITASELATG 600  
DB 541 RFGSSIASVRDLNODSYNDVVGAPLENDHAGALVIPHGRGSLTKTPKQITASELATG 600  
QY 601 LQYFCGCSIHGQDLNEDGLIDLAVGALGNVILMSRPVQINASLHPSPSKINIFHRDCK 660  
DB 601 LQYFCGCSIHGQDLNEDGLIDLAVGALGNVILMSRPVQINASLHPSPSKINIFHRDCK 660  
QY 661 RSGEDATCLAAFLCFTPIFLAPHEQTTTIGIRYNATMDERYTPRAHLDGEGDRTNRAV 720  
DB 661 RSGEDATCLAAFLCFTPIFLAPHEQTTTIGIRYNATMDERYTPRAHLDGEGDRTNRAV 720  
QY 721 LLSGQQLCERINFHVLDTDADYVKPVTFSVEYSLEDDPHGPMDDGWPTTLRSVPFWNG 780  
DB 721 LLSGQQLCERINFHVLDTDADYVKPVTFSVEYSLEDDPHGPMDDGWPTTLRSVPFWNG 780  
QY 781 CNEDEHCVDPDILVLDARSPLTAMEYICQVLRKPAQDCSAYTILSDTTVTTFIESTQRVAV 840  
DB 781 CNEDEHCVDPDILVLDARSPLTAMEYICQVLRKPAQDCSAYTILSDTTVTTFIESTQRVAV 840  
QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKESDGSIECVNERRQKQVCNVSYPFF 900  
DB 841 EATLENRGENAYSTVLNISQSANLQFASLIQKESDGSIECVNERRQKQVCNVSYPFF 900  
QY 901 RAKAKVAPRLDSEFSKSIPLHLEIELAAGSDNSNERDSTKEDNVAPLPHIKEYEADVLT 960  
DB 901 RAKAKVAPRLDSEFSKSIPLHLEIELAAGSDNSNERDSTKEDNVAPLPHIKEYEADVLT 960

QY 961 RSSSLSHYEVKLNSSLYERYDGIQPPSCIPRIOMGLFPIHGMKMTIPIATRSNRL 1020  
Db 961 RSSSLSHYEVKPNSSLYERYDGIQPPSCIPRIOMGLFPIHGMKMTIPIATRSNRL 1020  
QY 1021 KLRDLFTDEANTSCNIGWSTVRPTPVEEDLRAPQLNHSNDDVVSINCNRLVPOBI 1080  
Db 1021 KLRDLFTDEANTSCNIGWSTVRPTPVEEDLRAPQLNHSNDDVVSINCNRLVPOBI 1080  
QY 1081 NFHLGNLWLSLKALKYKSMKIMVNAALQORQFHSPPIFREEDPSQIIFESIKQBDWQV 1140  
Db 1081 NFHLGNLWLSLKALKYKSMKIMVNAALQORQFHSPPIFREEDPSQIIFESIKQBDWQV 1140  
QY 1141 PIWIIWGSTLGLLLALLVIALWLKLGFFPSARRRERGLDPTPKVLE 1188  
Db 1141 PIWIIWGSTLGLLLALLVIALWLKLGFFPSARRRERGLDPTPKVLE 1188

RESULT 5  
ABG12949  
ID ABG12949 standard; protein; 1189 AA.  
XX AC ABG12949;  
XX DT 13-FEB-2002 (first entry)  
XX DE Novel human diagnostic protein #12940.  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX OS Homo sapiens.  
XX PN WO200175067-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US008631.  
XX PR 31-MAR-2000; 2000US-00540217.  
XX PR 23-AUG-2000; 2000US-00649167.  
XX PA (HVS-) HVSQ INC.  
XX PI Dmanac RT, Liu C, Tang YT;  
XX WPI; 2001-639362/73.  
XX N-PSDB; AAS77136.  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
PS Claim 20; SEQ ID NO 43308; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1189 AA;  
Query Match 99.5%; Score 6192.5; DB 4; Length 1189;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1185; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
QY 1 MDLPRGLVAVALSILWPGFTDTFNNMTKPRVIFGSRTPAFGGYVQQHDIISGNKWLTVGA 60  
Db 1 MDLPRGLVAVALSILWPGFTDTFNNMTKPRVIFGSRTPAFGGYVQQHDIISGNKWLTVGA 60  
QY 61 PLEWNGYQKGTGVYKCPVINGNCTKLNLRGRVTLNVSERKDNMRGLGLSLATNPKDNSFLA 120  
Db 61 PLEWNGYQKGTGVYKCPVINGNCTKLNLRGRVTLNVSERKDNMRGLGLSLATNPKDNSFLA 120  
QY 121 CSPLWSHECGSSYYTTGCMRSVNSNFRFSTKVPALQRCQTYMDIVIVLDGNSIYPWVE 180  
Db 121 CSPLWSHECGSSYYTTGCMRSVNSNFRFSTKVPALQRCQTYMDIVIVLDGNSIYPWVE 180  
QY 181 VQHFLINILKKFYIGPGQIQGVVQYGEDVVEHFLNDYRSVKQVVEASHIEQGGTET 240  
Db 181 VQHFLINILKKFYIGPGQIQGVVQYGEDVVEHFLNDYRSVKQVVEASHIEQGGTET 240  
QY 241 RTAFGIEFARSEAFQKGGKGAKKVMIVITDGEHSDSPLEKVIQSSERDQVRYAVAVL 300  
Db 241 RTAFGIEFARSEAFQKGGKGAKKVMIVITDGEHSDSPLEKVIQSSERDQVRYAVAVL 300  
QY 301 GYNNRGINPETFLNEIKYIASDDDDGHPFVNTDEAALKDIDALGDRIFSLGNTKNMET 360  
Db 301 GYNNRGINPETFLNEIKYIASDDDDGHPFVNTDEAALKDIDALGDRIFSLGNTKNMET 360  
QY 361 SFGLEMSQTFSSHVVEDGVLGAVGAYDWMGAVLKETSGAKVIFLRYSYKEPPEELKN 420  
Db 361 SFGLEMSQTFSSHVVEDGVLGAVGAYDWMGAVLKETSGAKVIFLRYSYKEPPEELKN 420  
QY 421 HGAYLGYTVTSVSSRQGRVTVAGAPRNFHTKVLFTMHNNRSLTIHQMRGQOIGSYF 480  
Db 421 HGAYLGYTVTSVSSRQGRVTVAGAPRNFHTKVLFTMHNNRSLTIHQMRGQOIGSYF 480  
QY 481 GSEITSDIDGQVTDVLLVGAPMYFNEGRGRGVVYVELRQNRVYNGTLKDSYQNA 540  
Db 481 GSEITSDIDGQVTDVLLVGAPMYFNEGRGRGVVYVELRQNRVYNGTLKDSYQNA 540  
QY 541 RFGSSIASVRDLNQDSVNDVVGAPLEDNHAGAIYIFHGFRGSIILKTPKQITASELATG 600  
Db 541 RFGSSIASVRDLNQDSVNDVVGAPLEDNHAGAIYIFHGFRGSIILKTPKQITASELATG 600  
QY 601 LQYFGCSITHGOLDNEDGLIDLAVGALGNVILWSRPVQINASHLHFPESKINI FHRDCK 660  
Db 601 LQYFGCSITHGOLDNEDGLIDLAVGALGNVILWSRPVQINASHLHFPESKINI FHRDCK 660  
QY 661 RSGRDATCLAAFLCPTPIFLAPHFQTTTGVIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720  
Db 661 RSGRDATCLAAFLCPTPIFLAPHFQTTTGVIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720  
QY 721 LSSQRCICERINFHVLDTADYKVPVSFVSELDPDHGMKDDGWPTTLRVSVFPWNG 780  
Db 721 LSSQRCICERINFHVLDTADYKVPVSFVSELDPDHGMKDDGWPTTLRVSVFPWNG 780  
QY 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLSPDTTFFIIESTQRVAV 840  
Db 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLSPDTTFFIIESTQRVAV 840  
QY 841 EATLENRENAYSTVNLISQANLOFASLIQKESDGSIECNEERRLOKQVCNVSYPFF 900  
Db 841 EATLENRENAYSTVNLISQANLOFASLIQKESDGSIECNEERRLOKQVCNVSYPFF 900  
QY 901 RAKAKVAFELSDSPSKSIPLHLEITELAAAGSDNSNERDSTKEDNVAPLFHLYEADVLFT 960  
Db 901 RAKAKVAFELSDSPSKSIPLHLEITELAAAGSDNSNERDSTKEDNVAPLFHLYEADVLFT 960

Db 901 RAKAVARLDFEFKSIPLHLLEIELAGSDNSNERDSTKEDNVAPLRFHLKYEADVLT 960  
Qy 961 RSSLSHYVVKLNSSLRYDGGPPSCIFRIQNLGLFPIHGMXMKITPIATRSGRLL 1020  
Db 961 RSSLSHYVVKLNSSLRYDGGPPSCIFRIQNLGLFPIHGMXMKITPIATRSGRLL 1020  
Qy 1021 KLRDLFTDB-ANTSCNINWNSYETPTVREDLRAPOLNHSNDVVSINCNIRLVPQOE 1079  
Db 1021 KURDLFTDEVANTSCNINWNSYETPTVREDLRAPOLNHSNDVVSINCNIRLVPQOE 1080  
Qy 1080 INFHLGNLWLSLXALKYKSMKIMVNAALQRFHSPFIIFREEDPSROIETFSIKQEDWQ 1139  
Db 1081 INFHLGNLWLSLXALKYKSMKIMVNAALQRFHSPFIIFREEDPSROIETFSIKQEDWQ 1140  
Qy 1140 VPIWIIIVGSTGGLHLLLVLLALRKLGFFRSAREREPGLDPTPKVLE 1188  
Db 1141 VPIWIIIVGSTGGLHLLLVLLALWKLGFFRSAREREPGLDPTPKVLE 1189

RESULT 6  
AAU14467  
ID AAU14467 standard; protein; 1188 AA.  
XX AAU14467;  
AC AAU14467;  
DT 24-OCT-2001 (first entry)  
DE Human novel protein #338.  
DE Human; novel protein; Antianaemic; osteopathic; antiinflammatory;  
KW immunomodulatory; cycostatic; neuroprotective; vulnerary; nootropic;  
KW anticonvulsant; antiallergic; cerebroprotective; antifungal; antiviral;  
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;  
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
KW tissue regeneration; immune disorder.  
XX Homo sapiens.  
OS Homo sapiens.  
XX WO200155437-A2.  
PN 02-AUG-2001.  
PD 25-JAN-2001; 2001WO-US002623.  
PF 25-JAN-2000; 2000US-00491404.  
PR (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT;  
XX WPI; 2001-451939/48.  
XX N-PSDB; AAS22772.  
XX Isolated polypeptides useful for treating anti-inflammatory diseases,  
PT nervous system disorders, and for regenerating bone and cartilage.  
PT Example 4; Page 828-831; 894pp; English.  
PS The invention relates to polynucleotides encoding novel human proteins or  
CC their active domains. The polypeptides, polynucleotides and antibodies  
CC raised against the polypeptides are used in a method of treatment of a  
CC mammal and prevention of disorders caused by the aberrant protein  
CC expression or activity. The polypeptides can be used as molecular weight  
CC markers, food supplements, and in antibody production. The polypeptides  
CC are used to identify compounds which bind to the polypeptides.  
CC Polynucleotides of the invention are used as probes and primers, for  
CC sequencing, for chromosome or gene mapping, in the production of  
CC recombinant proteins, and in generating anti-sense DNA or RNA and in gene  
CC therapy. Polypeptides of the invention can be used to target drugs to a  
CC tumour, in assays to determine biological activity, to raise  
CC antibodies/elicit an immune response, to determine quantitative protein  
CC levels, as tissue markers, and to isolate receptors or ligands.

CC Polypeptides of the invention may also be useful in treating platelet  
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
CC the proliferation, differentiation and survival of stem cells, as a  
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
CC fungal infection or from autoimmunity, cancer, allergy, asthma, graft-  
CC versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory  
CC diseases, nervous system disorders, and infection. The present sequence  
CC represents a protein of the invention  
XX Sequence 1188 AA;  
SQ Query Match 99.5%; Score 6192; DS 4; Length 1188;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1183; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 MDLPRLGLVAVWALSLLWPGFTDTFNMDTRKPRVIPSRTAPFGYTVQOQHDISGNKWLWVGA 60  
Db 1 MDLPRLGLVAVWALSLLWPGFTDTFNMDTRKPRVIPSRTAPFGYTVQOQHDISGNKWLWVGA 60  
Qy 61 PLETNGYQKTGDVYKCPVIHGNCTKLNLRVTLSNVSEKDNVRLGLSLATNPXNSFLA 120  
Db 61 PLETNGYQKTGDVYKCPVIHGNCTKLNLRVTLSNVSEKDNVRLGLSLATNPXNSFLA 120  
Qy 121 CSPLWSHECGSSVYVTCMCSEKSNRFRFSKTVAPALQRCOTYMDIVIVLDSGNSIYPWVE 180  
Db 121 CSPLWSHECGSSVYVTCMCSEKSNRFRFSKTVAPALQRCOTYMDIVIVLDSGNSIYPWVE 180  
Qy 181 VQFLINILKFKYIGPQIQVGVVQYGEDVVEPHLDYRSVKDVVEAASHIEBQGTET 240  
Db 181 VQFLINILKFKYIGPQIQVGVVQYGEDVVEPHLDYRSVKDVVEAASHIEBQGTET 240  
Qy 241 RTAFGIEFARSEAFQKGRGAKKMWIVITDGHSDSPDLKVKVQQSERDNVTVAVAVL 300  
Db 241 RTAFGIEFARSEAFQKGRGAKKMWIVITDGHSDSPDLKVKVQQSERDNVTVAVAVL 300  
Qy 301 GYNRGINPETFLNEIKYIASDPDDKHFNVTDEAALKDIDVADGDRIFSEGTNKNET 360  
Db 301 GYNRGINPETFLNEIKYIASDPDDKHFNVTDEAALKDIDVADGDRIFSEGTNKNET 360  
Qy 361 SFGLEMSQTGFSSHVVEDGVLGNVGNAYDNGAVLAKETSAGKVIPLRESYLKEPBEUKN 420  
Db 361 SFGLEMSQTGFSSHVVEDGVLGNVGNAYDNGAVLAKETSAGKVIPLRESYLKEPBEUKN 420  
Qy 421 HGAYLGYTVTSVSSRQGRVYVAGAPRFNETGKVIPLFMENRSLTHQAMRGQIGSYF 480  
Db 421 HGAYLGYTVTSVSSRQGRVYVAGAPRFNETGKVIPLFMENRSLTHQAMRGQIGSYF 480  
Qy 481 GSEITSDVIDDGDVTVLLVAGAPMYFNEGRERKQVYVELRQNFVYNGTLKDSHSYQNA 540  
Db 481 GSEITSDVIDDGDVTVLLVAGAPMYFNEGRERKQVYVELRQNFVYNGTLKDSHSYQNA 540  
Qy 541 RFGSSIASVRDLNODSVNDVVVCGAPLEDNHAGAIYIPHGFRGSLTKPKQITASELATG 600  
Db 541 RFGSSIASVRDLNODSVNDVVVCGAPLEDNHAGAIYIPHGFRGSLTKPKQITASELATG 600  
Qy 601 LQVFGCSIHQOLDLNEBGLDIDLVAGALGNVILWSRPVVOINASLHFPESKIMIFERDCK 660  
Db 601 LQVFGCSIHQOLDLNEBGLDIDLVAGALGNVILWSRPVVOINASLHFPESKIMIFERDCK 660  
Qy 661 RSRGDATCLAAFLCFTPIPLAHPFQTTVGIRYNATMDERRYPRAHLDGGDRFTNRVAV 720  
Db 661 RSRGDATCLAAFLCFTPIPLAHPFQTTVGIRYNATMDERRYPRAHLDGGDRFTNRVAV 720  
Qy 721 LLSGGQELCERINFHVLDTADYVVKPVTFSVEYSLEDDPHGFMLDGQWPTTLRVSPFWNG 780  
Db 721 LLSGGQELCERINFHVLDTADYVVKPVTFSVEYSLEDDPHGFMLDGQWPTTLRVSPFWNG 780  
Qy 781 CNEDEHCVPDLVLDARSJLPTAMEYCORVLRKPAQDCSAVTLISFDITVFIIESTRORVAV 840  
Db 781 CNEDEHCVPDLVLDARSJLPTAMEYCORVLRKPAQDCSAVTLISFDITVFIIESTRORVAV 840



QY 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIKEDSIECVNEERRLQKQVCNVSYPFF 900  
 DB 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIKEDSIECVNEERRLQKQVCNVSYPFF 900  
 QY 901 RAKAVAFRLDSEPSKSIPLHLLEIELAAGSDNERDSTKEDNAPLRFHLKYADVLFT 960  
 DB 901 RAKAVAFRLDSEPSKSIPLHLLEIELAAGSDNERDSTKEDNAPLRFHLKYADVLFT 960  
 QY 961 RSSLSHYEVKLNLSLERYDGIQPPFSCIFRIQNLGLPPIHGMWMTIPIATRSNRL 1020  
 DB 961 RSSLSHYEVKLNLSLERYDGIQPPFSCIFRIQNLGLPPIHGMWMTIPIATRSNRL 1020  
 QY 1021 KLDFLTDEANTSCINWNSSTYRTPVEEDLERAPOLNHSNSDVVSNINRLVFNQEI 1080  
 DB 1021 KLDFLTDEANTSCINWNSSTYRTPVEEDLERAPOLNHSNSDVVSNINRLVFNQEI 1080  
 QY 1081 NFHLGLNLWLSRLKALKYKSMIMVNAALQROFHPFIFREEDPSRQIVFELSKQEDMQV 1140  
 DB 1081 NFHLGLNLWLSRLKALKYKSMIMVNAALQROFHPFIFREEDPSRQIVFELSKQEDMQV 1140  
 QY 1141 PIWIIIVGSTLGLLALLVLAALKLGFRRSARRRREPGLDTPPKVLE 1188  
 DB 1141 PIWIIIVGSTLGLLALLVLAALKLGFRRSARRRREPGLDTPPKVLE 1188  
 RESULT 7  
 ADE09956  
 ID ADE09956 standard; protein; 1188 AA.  
 AC ADE09956;  
 DT 29-JAN-2004 (first entry)  
 XX Novel protein-related contig polypeptide sequence #544.  
 XX novel gene; novel protein; tissue marker; molecular weight marker;  
 KW chromosome marker; Genetic disorder; contig.  
 XX Unidentified.  
 XX WO2003054152-A2.  
 XX 03-JUL-2003.  
 XX 10-DEC-2002; 2002WO-US039555.  
 XX 10-DEC-2001; 2001US-0339739P.  
 PR 11-DEC-2001; 2001US-0339453P.  
 PR 14-MAR-2002; 2002US-0365091P.  
 PR 14-MAR-2002; 2002US-0365384P.  
 PR 12-APR-2002; 2002US-0372381P.  
 PR 12-APR-2002; 2002US-0372615P.  
 PR 22-APR-2002; 2002US-00128558.  
 PR 24-APR-2002; 2002US-0376045P.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;  
 PI Chosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;  
 PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;  
 XX WPI; 2003-569235/53.  
 XX New polynucleotides, useful for expressing recombinant proteins for  
 PT analysis, characterization or therapeutic use, or as markers for tissues  
 PT in which the corresponding protein is preferentially expressed.  
 XX  
 PS Disclosure; SEQ ID NO 3022; 1177bp; English.  
 CC The invention comprises the amino acid and coding sequences of novel  
 CC proteins. The DNA and protein sequences of the invention are useful as:  
 CC markers for tissues in which the corresponding protein is preferentially

CC expressed; as molecular weight markers on gels; as chromosome markers or  
 CC tags; to identify chromosomes or to map related gene positions; and to  
 CC compare with endogenous DNA sequences in patients to identify potential  
 CC genetic disorders. The present amino acid sequence was used in the  
 CC exemplification of the invention.  
 XX  
 SQ Sequence 1188 AA;  
 Query Match 99.5%; Score 6192; DB 7; Length 1188;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 1183; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MDLPRGLVAVAWLSLWPGTDTFNMDTRKPRVTPGSRAPFGYVQOHDISGNKWLVGA 60  
 DB 1 MDLPRGLVAVAWLSLWPGTDTFNMDTRKPRVTPGSRAPFGYVQOHDISGNKWLVGA 60  
 QY 61 PLENGYQKTDVYKCFVIHGNCTKLNLGRVTLNSVSRKDNMRGLGLSLATNPXNSFLA 120  
 DB 61 PLENGYQKTDVYKCFVIHGNCTKLNLGRVTLNSVSRKDNMRGLGLSLATNPXNSFLA 120  
 QY 121 CSPLWSEHCSSYTTGMCVRVNSNFRFSTKVPALQRCOTYMDIVIVLQGSNSIYPWE 180  
 DB 121 CSPLWSEHCSSYTTGMCVRVNSNFRFSTKVPALQRCOTYMDIVIVLQGSNSIYPWE 180  
 QY 181 VQFLINILKKFYIGPGQIQGVVQYGEDVVFHFLNDYRSVKDWVAAASHIEQGGTET 240  
 DB 181 VQFLINILKKFYIGPGQIQGVVQYGEDVVFHFLNDYRSVKDWVAAASHIEQGGTET 240  
 QY 241 RTAFGIEFARSEAPQKGRGAKKVMIVITDGHSDSPDLKVIQOESRDNTRYAVAVL 300  
 DB 241 RTAFGIEFARSEAPQKGRGAKKVMIVITDGHSDSPDLKVIQOESRDNTRYAVAVL 300  
 QY 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNFVITDEAALKDVIDALGDRIIFSLEGTNKNET 360  
 DB 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNFVITDEAALKDVIDALGDRIIFSLEGTNKNET 360  
 QY 361 SFGLEMSQTFSSHVEDGVLLGAVGAYDNGAVLKTSGKVIPLRESYLKEPPELKN 420  
 DB 361 SFGLEMSQTFSSHVEDGVLLGAVGAYDNGAVLKTSGKVIPLRESYLKEPPELKN 420  
 QY 421 HGAYLGVTVTSVSSRQGRVYVAGAPFNHTGKVLFTMNNSRLTIHQAMRGQIQISYF 480  
 DB 421 HGAYLGVTVTSVSSRQGRVYVAGAPFNHTGKVLFTMNNSRLTIHQAMRGQIQISYF 480  
 QY 481 GSEITSDVIDDGVTVLLVGA PMYFNEGRERGVYVYELRQNRPFVNGTLKDSHSYQNA 540  
 DB 481 GSEITSDVIDDGVTVLLVGA PMYFNEGRERGVYVYELRQNRPFVNGTLKDSHSYQNA 540  
 QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGRGSLTKTPKORITASELATG 600  
 DB 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGRGSLTKTPKORITASELATG 600  
 QY 601 LQYFGCSIHQQLDNLNEDGLIDLAVGALGNNAVILMSRPVQVQINASLHFEPSKINIFHRDCK 660  
 DB 601 LQYFGCSIHQQLDNLNEDGLIDLAVGALGNNAVILMSRPVQVQINASLHFEPSKINIFHRDCK 660  
 QY 661 RSGRDATAFLAFLCTPIELAPHFQTTVGIRYNATMDERRYPRAHLDEGGDRFTNRVAV 720  
 DB 661 RSGRDATAFLAFLCTPIELAPHFQTTVGIRYNATMDERRYPRAHLDEGGDRFTNRVAV 720  
 QY 721 LLSGQELCERINFHVLDTADYVKPVTFSVEYSLEDPDHGMPLDDGPTTLRVSPFWNG 780  
 DB 721 LLSGQELCERINFHVLDTADYVKPVTFSVEYSLEDPDHGMPLDDGPTTLRVSPFWNG 780  
 QY 781 CNEDEHCVPOLVLDARSDDLPTAMEYCORVLKPAQDCSAYTSLSDTTVFIIESTRQAVAV 840  
 DB 781 CNEDEHCVPOLVLDARSDDLPTAMEYCORVLKPAQDCSAYTSLSDTTVFIIESTRQAVAV 840  
 QY 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIKEDSIECVNEERRLQKQVCNVSYPFF 900  
 DB 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIKEDSIECVNEERRLQKQVCNVSYPFF 900  
 QY 901 RAKAVAFRLDSEPSKSIPLHLLEIELAAGSDNERDSTKEDNAPLRFHLKYADVLFT 960

Db 901 RAKAKVAFRLDFEFKSIPLHLELELAAGSDNSRDKENVNAPLRFHLKYEVDVLT 960  
QY 961 RSSLSHYEVKLNSSLERYDGIQPPFSCIFRIQNLGLPPIHGMKKIPIATRSNRL 1020  
Db 961 RSSLSHYEVKLNSSLERYDGIQPPFSCIFRIQNLGLPPIHGMKKIPIATRSNRL 1020  
QY 1021 KURDLTDEANTSCNMGWNSTEYRTPVEEDLRRAPQLNHSNDVVSNCNIRLVPNQEI 1080  
Db 1021 KURDLTDEANTSCNMGWNSTEYRTPVEEDLRRAPQLNHSNDVVSNCNIRLVPNQEI 1080  
QY 1081 NFHLGNLWLRSLKALKYKSMIMVNAALQRFHSPFFIREDPSRQIEFEISKQEDQV 1140  
Db 1081 NFHLGNLWLRSLKALKYKSMIMVNAALQRFHSPFFIREDPSRQIEFEISKQEDQV 1140  
QY 1141 PIWIIVGSTLGGILLALLVIALKLGFFRSARRRRPGLDTPKYLE 1188  
Db 1141 PIWIIVGSTLGGILLALLVIALKLGFFRSARRRRPGLDTPKYLE 1188

RESULT 8  
ID AAB25582  
XX AAB25582 standard; protein; 1189 AA.  
AC AAB25582;  
DT 21-NOV-2000 (first entry)  
XX  
DE ITGall protein encoded by human secreted protein gene #7.  
XX  
KW Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;  
KW antirheumatic; dermatological; antiproliferative; antiarteriosclerotic;  
KW anticancer; vulnery; antiviral; antibacterial; antifungal;  
KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;  
KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;  
KW Crohn's disease; nephritis; hyperproliferative disorder;  
KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;  
KW melanoma; lymphoma; wound healing; human; chromosome 15 q22.3-23.  
XX  
OS Homo sapiens.  
XX WO200029435-A1.  
PN 25-MAY-2000.  
PD  
XX 27-OCT-1999; 99WO-US025031.  
PF  
XX 28-OCT-1998; 98US-0105971P.  
PR  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Ni J, Ruben SM, Olsen HS, Young PE, Kenny JU, Moore PA, Wei Y;  
PI Greene JM;  
PI  
XX WPI: 2000-387742/33.  
DR N-PSDB; AAB80612.  
XX  
XX Isolated nucleic acid molecules encoding human secreted proteins are used  
PT for the prevention, amelioration and treatment of autoimmune,  
PT inflammatory, hyperproliferative and cardiovascular disorders, cancer,  
PT wounds, and infectious diseases.  
PT  
PS Claim 1: Fig 19A-F; 803pp; English.

CC The present invention relates to 12 secreted human proteins and the  
CC nucleotide sequences encoding them. The polynucleotide sequences given in  
CC AAA80606-A80623 encode the 12 secreted protein sequences given in  
CC AAB25576-B25593. The human secreted proteins have various activities  
CC dependent on the tissues in which they are expressed. Examples of the  
CC activities of the proteins include: immunosuppressant; anti-inflammatory;  
CC antirheumatic; dermatological; antiproliferative; anti-infectious;  
CC antitumor; anticancer; vulnery; antiviral; antibacterial;  
CC and antifungal activity. The proteins, polypeptides, agonists and

CC antagonists may be used to treat prevent and/or diagnose various disease,  
CC disorders and conditions examples of which include: immune disorders e.g.  
CC Addison's disease, rheumatoid arthritis, dermatitis, and multiple  
CC sclerosis; inflammatory disorders e.g. inflammatory bowel disease,  
CC Crohn's disease and nephritis; hyperproliferative disorders such as  
CC paraproteinemia and purpura; cardiovascular disorders e.g. coronary  
CC arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The  
CC proteins and polynucleotide sequences may also be used in wound healing  
CC and the treatment of infectious diseases. The human secreted protein gene  
CC #7 and protein sequences are represented in sequences AAB80612 and  
CC AAB25582. Secreted protein gene #7 is located at position chromosome 15  
CC q22.3-23. Sequences AAB80652-A80661 represent genes which are related to  
CC the secreted protein gene#7  
XX  
SQ Sequence 1189 AA,

Query Match 99.4%; Score 6188.5; DB 3; Length 1189;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1184; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
QY 1 MDLPRGLVAVWALSILWPGFTDTFNMDTRKPRVIPSRTAPFGYVQOHDISGNKWLWVGA 60  
Db 1 MDLPRGLVAVWALSILWPGFTDTFNMDTRKPRVIPSRTAPFGYVQOHDISGNKWLWVGA 60  
QY 61 PLETNGYQKTGDVYKCFVIEGNCCTKLNLGRVTLNVSERKDNMRGLSLATNPKDNSFLA 120  
Db 61 PLETNGYQKTGDVYKCFVIEGNCCTKLNLGRVTLNVSERKDNMRGLSLATNPKDNSFLA 120  
QY 121 CSPLWSHECSSYYTTCMCSSVNSNPRFSKTVAPALQRCQYMDIVIVLDCGNSNLYPWVE 180  
Db 121 CSPLWSHECSSYYTTCMCSSVNSNPRFSKTVAPALQRCQYMDIVIVLDCGNSNLYPWVE 180  
QY 181 VQFELNLIKFFYIGPGQIQGVVQYGEDVVBHFLANDYRSVKDVEAASHIEBGGTET 240  
Db 181 VQFELNLIKFFYIGPGQIQGVVQYGEDVVBHFLANDYRSVKDVEAASHIEBGGTET 240  
QY 241 RTAGIFPAREARQKGRKAKKQVIMVITDGHSDSPDLKVKQOSRDNVTVYAVAVL 300  
Db 241 RTAGIFPAREARQKGRKAKKQVIMVITDGHSDSPDLKVKQOSRDNVTVYAVAVL 300  
QY 301 GYNERGINPETFLNEIKYIASDPDDKHFFNVTDAAALKOIVDALGDRIFSLGNTKNET 360  
Db 301 GYNERGINPETFLNEIKYIASDPDDKHFFNVTDAAALKOIVDALGDRIFSLGNTKNET 360  
QY 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPEELKN 420  
Db 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPEELKN 420  
QY 421 HGAVLGYVTWSSVSSROGRVYVAGAPRPNHTGKVIPLFMHNNRSLTTHQAMRGQIGSYF 480  
Db 421 HGAVLGYVTWSSVSSROGRVYVAGAPRPNHTGKVIPLFMHNNRSLTTHQAMRGQIGSYF 480  
QY 481 GSEITSVDIDGDGVTDLVLCAPMYFNEGRERKGVYVYELRQNRVFNVTGLKDSHSYQNA 540  
Db 481 GSEITSVDIDGDGVTDLVLCAPMYFNEGRERKGVYVYELRQNRVFNVTGLKDSHSYQNA 540  
QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGPFGSILKTPKQITASELATG 600  
Db 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGPFGSILKTPKQITASELATG 600  
QY 601 LQYFGCSIHGOLDLNEDELIDVAVGALGNVILWSRPVQVQINASLHFEPSKINIFHRDCK 660  
Db 601 LQYFGCSIHGOLDLNEDELIDVAVGALGNVILWSRPVQVQINASLHFEPSKINIFHRDCK 660  
QY 661 RSGDADYCLAAFLCTFIFLAPHPTTTVGRYRNATWDEREYTPRAHLDEGDRFTNAV 720  
Db 661 RSGDADYCLAAFLCTFIFLAPHPTTTVGRYRNATWDEREYTPRAHLDEGDRFTNAV 720  
QY 721 LLSGQELCERINPHVLDADYVKPVTFVSVEYSLEDDHGMPLDDGMPPTTLRVSVFPWNG 780  
Db 721 LLSGQELCERINPHVLDADYVKPVTFVSVEYSLEDDHGMPLDDGMPPTTLRVSVFPWNG 780  
QY 781 CNEDEHCVDPDLVLDARSDDLPTAMEYCORVLKAKQDCASTLSDFTTPTTETSTQRVAV 840

Db 781 CNEDEHCPDLVDARSDDLPATMEYCORVLRPAQDCSAYTUSFDVTTFIISTQRVAV 840  
Qy 841 EATLENGENAYSTVLNISOANLQFASLIQKEDSDGSIQECVNEERLQKQVNSYPPF 900  
Db 841 EATLENGENAYSTVLNISOANLQFASLIQKEDSDGSIQECVNEERLQKQVNSYPPF 900  
Qy 901 RAKAKVAFRLDSFSSKIFLHLEIELAAGSDSNERDSTKEDNAPLRPHLYEADVLPT 960  
Db 901 RAKAKVAFRLDSFSSKIFLHLEIELAAGSDSNERDSTKEDNAPLRPHLYEADVLPT 960  
Qy 961 RSSLSHYEVKNSLERYDGIQPPFSCIFRIQNLGLFPIHGMKTIPIATRSNRL 1020  
Db 961 RSSLSHYEVKNSLERYDGIQPPFSCIFRIQNLGLFPIHGMKTIPIATRSNRL 1020  
Qy 1021 KLBDPLTDE-ANTSCINWGNSTYRTPVEEDLRRAPQLNHSNDVSNIRLVPNOE 1079  
Db 1021 KLBDPLTDE-ANTSCINWGNSTYRTPVEEDLRRAPQLNHSNDVSNIRLVPNOE 1080  
Qy 1080 INFHLLGNLWLRSLKALKYKSMKIMVNAALQROFHSPFFREEDPSRQIEFEISKQEDWQ 1139  
Db 1081 INFHLLGNLWLRSLKALKYKSMKIMVNAALQROFHSPFFREEDPSRQIEFEISKQEDWQ 1140  
Qy 1140 VPIWIIIVGSTLGGILLIALLVLALRKLGFPRSRARRRREPGLDPTPKVLE 1188  
Db 1141 VPIWIIIVGSTLGGILLIALLVLALRKLGFPRSRARRRREPGLDPTPKVLE 1189

RESULT 9  
ABR58364 standard; protein; 1189 AA.  
ID ABR58364;  
XX AC ABR58364;  
XX DT 07-JUL-2003 (first entry)  
XX DE Human NOV2a.  
XX KW Human; NOV; antidiabetic; anorectic; antibacterial; virucide;  
KW immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia;  
KW antiparkinsonian; antilipaeamic; gene therapy; metabolic disorder;  
KW diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;  
KW neurodegenerative disorder; Alzheimer's disease; immune disorder;  
KW haematopoietic disorder.  
XX OS Homo sapiens.  
XX PN WO2003029423-A2.  
XX PD 10-APR-2003.  
XX PF 02-OCT-2002; 2002WO-US031358.  
XX PR 02-OCT-2001; 2001US-0326483P.  
PR 05-OCT-2001; 2001US-0327342P.  
PR 09-OCT-2001; 2001US-0327917P.  
PR 09-OCT-2001; 2001US-0328029P.  
PR 09-OCT-2001; 2001US-0328044P.  
PR 09-OCT-2001; 2001US-0328056P.  
PR 12-OCT-2001; 2001US-0328849P.  
PR 15-OCT-2001; 2001US-0329414P.  
PR 17-OCT-2001; 2001US-0330142P.  
PR 22-OCT-2001; 2001US-0341058P.  
PR 24-OCT-2001; 2001US-0339266P.  
PR 29-OCT-2001; 2001US-0343629P.  
PR 29-OCT-2001; 2001US-0349575P.  
PR 01-NOV-2001; 2001US-0346337P.  
PR 12-APR-2002; 2002US-0371972P.  
PR 12-APR-2002; 2002US-0371980P.  
PR 17-APR-2002; 2002US-0373261P.  
PR 19-APR-2002; 2002US-0373805P.  
PR 23-APR-2002; 2002US-0374738P.  
PR 16-MAY-2002; 2002US-0381101P.

PR 17-MAY-2002; 2002US-0381635P.  
PR 29-MAY-2002; 2002US-0383830P.  
PR 01-OCT-2002; 2002US-00262839.  
XX (CURA-) CURAGEN CORP.  
XX PI Alsbrook JP, Ellerman DW, Boldog PL, Burgess CE, Catterton E;  
PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;  
PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;  
PI Rothenberg ME, Shinkens RA, Smithson G, Spytek KA, Taupier RJ;  
PI Vernet CM, Voss EZ, Zerhusen BD, Zhong M;  
XX WPI; 2003-381625/36.  
XX DR N-PSDB; ACC72076.  
XX PT NOVX polypeptides and nucleic acids useful for diagnosing, preventing or  
PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or  
PT dyslipidemia, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX PS Claim 1; Page 105; 487pp; English.  
XX CC The present invention relates to novel human NOV proteins and their  
CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV  
CC proteins are useful in manufacturing a medicament for treating a syndrome  
CC associated with a human disease. The NOV proteins and coding sequences  
CC may be used to diagnose, treat or prevent metabolic disorders such as  
CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative  
CC disorders such as Alzheimer's disease or Parkinson's disease, immune  
CC disorders, haematopoietic disorders and various dyslipidaemias  
XX SQ Sequence 1189 AA;  
Query Match 99.4%; Score 6188.5; DB 6; Length 1189;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1184; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
Qy 1 MDLPGLVAVALSILWPGFTDFNMOTKPRVIPCGRRTAFGTYTQOHDISGNKVLVGA 60  
Db 1 MDLPGLVAVALSILWPGFTDFNMOTKPRVIPCGRRTAFGTYTQOHDISGNKVLVGA 60  
Qy 61 PLETTNGYQKTVGVVYKCPVHGNCTKLNLRVTVLSNVSRKONMRLGLSLATPKDNSFLA 120  
Db 61 PLETTNGYQKTVGVVYKCPVHGNCTKLNLRVTVLSNVSRKONMRLGLSLATPKDNSFLA 120  
Qy 121 CSPLMSHECGSSYYTTGMSRVNSNFRSKTVAPALQRCQTYMDIVIVLDSNSIYPWVE 180  
Db 121 CSPLMSHECGSSYYTTGMSRVNSNFRSKTVAPALQRCQTYMDIVIVLDSNSIYPWVE 180  
Qy 181 VOHFLINILKFPYIGPGQIQGVQVQYGEDVHVEFHLNDYRSVKDVVEAASHIEQGGTET 240  
Db 181 VOHFLINILKFPYIGPGQIQGVQVQYGEDVHVEFHLNDYRSVKDVVEAASHIEQGGTET 240  
Qy 241 RTAPGIEFARSAFQKGRKGAKKVMIVITDGEHSDSPDLKVIQOOSRDNTRYAVAVL 300  
Db 241 RTAPGIEFARSAFQKGRKGAKKVMIVITDGEHSDSPDLKVIQOOSRDNTRYAVAVL 300  
Qy 301 GYNNRGINPETFLNEIKYIASDDPKHFNVTDEAALKDIIVDALGDRIFSLGNTKNET 360  
Db 301 GYNNRGINPETFLNEIKYIASDDPKHFNVTDEAALKDIIVDALGDRIFSLGNTKNET 360  
Qy 361 SFGLEMSQTGFSSHVVDEGVLLGAYDNGAVLKETTSAGKVIPLRSYLKFPPELKN 420  
Db 361 SFGLEMSQTGFSSHVVDEGVLLGAYDNGAVLKETTSAGKVIPLRSYLKFPPELKN 420  
Qy 421 HGAYLGYTVTSVSSRQGEVYVAGAPRNFHTGKVLFTMHNRSLSLTHQAMGQQIGSYF 480  
Db 421 HGAYLGYTVTSVSSRQGEVYVAGAPRNFHTGKVLFTMHNRSLSLTHQAMGQQIGSYF 480  
Qy 481 GSEITTSVIDIDGQVTDVLLVAGAPMYFNEGRGRGVVYVELRQNRFFVYNGTLKDSHSYQNA 540  
Db 481 GSEITTSVIDIDGQVTDVLLVAGAPMYFNEGRGRGVVYVELRQNRFFVYNGTLKDSHSYQNA 540

QY 541 RFGSSIASVRDLNODSYNDVWVVGAPLEDNHAGAIYIFHGFRGSILKTKPKRITASELATG 600  
DB 541 RFGSSIASVRDLNODSYNDVWVVGAPLEDNHAGAIYIFHGFRGSILKTKPKRITASELATG 600  
QY 601 LOYFGCSIHGQDLNEDGLIDLVANGALGNVILMSRPVVOINASHLPSPSKINIFHRDCK 660  
DB 601 LOYFGCSIHGQDLNEDGLIDLVANGALGNVILMSRPVVOINASHLPSPSKINIFHRDCK 660  
QY 661 RSGRDATCLAAFLCTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDEGDRFTNRVAV 720  
DB 661 RSGRDATCLAAFLCTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDEGDRFTNRVAV 720  
QY 721 LLSQOELCERINFVLDATADYVVKVTFVSVEISLEDPDHGMPLDGDWPTTLRVSPFPWNG 780  
DB 721 LLSQOELCERINFVLDATADYVVKVTFVSVEISLEDPDHGMPLDGDWPTTLRVSPFPWNG 780  
QY 781 CNEDEHCVDPDLVDARSOLPTAMEYCORVLRKPAQDCSAYTILSPFTTVPFIIESRQVAV 840  
DB 781 CNEDEHCVDPDLVDARSOLPTAMEYCORVLRKPAQDCSAYTILSPFTTVPFIIESRQVAV 840  
QY 841 EATLENRGENAVSTVLNISQSANLOFASLIQKEDSDGSIQECVNEERRLOKQVCNVSYPFF 900  
DB 841 EATLENRGENAVSTVLNISQSANLOFASLIQKEDSDGSIQECVNEERRLOKQVCNVSYPFF 900  
QY 901 RAKAVAPFLDSEFSKSIPLHLELELAGSDNSRSDTKEDNVAPLRFHLKYRADVLFT 960  
DB 901 RAKAVAPFLDSEFSKSIPLHLELELAGSDNSRSDTKEDNVAPLRFHLKYRADVLFT 960  
QY 961 RSSLSHVEYKLNSSLRYDVGIPPPSCIFRIQNEGLPPIHGMKMTIPIATRSGNELL 1020  
DB 961 RSSLSHVEYKLNSSLRYDVGIPPPSCIFRIQNEGLPPIHGMKMTIPIATRSGNELL 1020  
QY 1021 KLKDLFTDE-ANTSCNINWGNSTVEYRTPVEEDLRAPQLNHSNDSVWVINCINIRLVNQE 1079  
DB 1021 KLKDLFTDE-ANTSCNINWGNSTVEYRTPVEEDLRAPQLNHSNDSVWVINCINIRLVNQE 1080  
QY 1080 INFHLLGNLWLSLKALKYKSKIMVNAALQRFHSPIFREEDPSROIPEISKQEDWQ 1139  
DB 1081 INFHLLGNLWLSLKALKYKSKIMVNAALQRFHSPIFREEDPSROIPEISKQEDWQ 1140  
QY 1140 VPIWIIIVSGTGLGLLLALLVLALXKLGFFRSARRRRRPPGLDPTPKVLE 1188  
DB 1141 VPIWIIIVSGTGLGLLLALLVLALXKLGFFRSARRRRRPPGLDPTPKVLE 1189

RESULT 10  
ADA27054  
ID ADA27054 standard; protein; 1189 AA.  
XX AC ADA27054;  
XX DT 20-NOV-2003 (first entry)  
XX DE Human novel secreted protein from cDNA HOHBV69 #1.  
XX KW cytosolic; antiinflammatory; immunomodulator; neuroprotective;  
KW hemostatic; gene therapy; cancer; inflammation; immune disorder;  
KW neurological disorder; blood clotting disorder; food additive;  
KW preservative; human; secreted protein.  
XX OS Homo sapiens.  
XX PN US2003055231-A1.  
XX PD 20-MAR-2003.  
XX PF 29-OCT-2001; 2001US-00984130.  
XX PR 28-OCT-1998; 98US-0105971P.  
XX PR 27-OCT-1999; 99WO-US025031.  
XX PR 19-APR-2000; 2000US-0198407P.  
XX PR 30-OCT-2000; 2000US-0243792P.  
XX PR 18-APR-2001; 2001US-00836353.

XX (NIJJ/) NI J.  
PA (YOUN/) YOUNG P E.  
PA (KERN/) KENNY J J.  
PA (OLSE/) OLSEN H S.  
PA (MOOR/) MOORE P A.  
PA (WEIY/) WEI Y.  
PA (GREE/) GREENE J M.  
PA (RUBE/) RUBEN S M.  
PA (LIUD/) LIU D.  
PA (CROC/) CROCKER P R.  
XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;  
PI Ruben SM, Liu D, Crocker PR;  
XX WPI; 2003-567103/53.  
DR N-PSDB; ADA27036.  
XX New human secreted nucleic acid molecules and polypeptides, useful for  
PT preventing, treating, or ameliorating a medical condition, such as  
PT cancer, inflammation, immune disorders, neurological and blood clotting  
PT disorders.  
XX Claim 11; Fig 19; 454pp; English.  
XX The invention relates to an isolated nucleic molecule that is at least  
CC 95% identical to 18 human cDNA sequences representing 12 novel genes  
CC encoding secreted proteins or a polynucleotide fragment of the cDNA  
CC sequence contained in American Type Culture Collection (ATCC) deposit No.  
CC defined in the specification, its species homologue, a variant or allelic  
CC variant of the polynucleotide having a polynucleotide capable of  
CC hybridising under conditions the polynucleotide, where the polynucleotide  
CC does not hybridise under stringent conditions to a nucleic acid molecule  
CC having a nucleotide sequence of only A or T residues. Also included are  
CC recombinant vectors, host cells (for producing the polypeptide), the  
CC secreted polypeptide (comprising a sequence that is at least 95%  
CC identical to a polypeptide fragment, domain, epitope, full-length  
CC protein, variant, allelic variant or species homologue), antibodies that  
CC specifically bind to the polypeptides, (diagnosing, treating, preventing  
CC or ameliorating a medical condition by administering the polynucleotide  
CC or the polypeptide, the gene corresponding to the cDNA sequence and  
CC identifying an activity in a biological assay (by expressing the cDNA  
CC sequence in a cell, isolating the supernatant, and detecting an activity  
CC in a biological assay and identifying the protein in the supernatant  
CC having the activity). The polypeptides, nucleic acids and antibodies are  
CC useful for diagnosing a pathological condition or a susceptibility to a  
CC pathological condition, for preventing, treating, or ameliorating a  
CC medical condition, such as cancer, inflammation and other immune  
CC disorders, neurological and blood clotting disorders (many examples are  
CC given in the specification). The nucleic acids are also useful for  
CC chromosome identification, radiation hybrid mapping or long-range  
CC restriction mapping. The polypeptides and antibodies are useful for  
CC providing immunological probes for differential identification of the  
CC tissues immunohistochemistry assays. The polypeptide, polynucleotide,  
CC agonist or antagonist may also be used as a food additive or preservative  
CC to increase or decrease storage capabilities, fat content or other  
CC nutritional components. The present is a secreted protein of the  
XX invention.  
XX Sequence 1189 AA;  
SQ

Query Match 99.4%; Score 6188.5; DB 6; Length 1189;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1184; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
QY 1 MDLPRGLVVAWALSILWPGFTDTFNMDTKPRVIPSRTAPFGYTVQQHDISEGNKLVVGA 60  
DB 1 MDLPRGLVVAWALSILWPGFTDTFNMDTKPRVIPSRTAPFGYTVQQHDISEGNKLVVGA 60  
QY 61 PLETNGYQKTGDYVKCPVTHGNCTKLNLRVTLSNVSEKKNMELGLSLATNPKDNSPLA 120  
DB 61 PLETNGYQKTGDYVKCPVTHGNCTKLNLRVTLSNVSEKKNMELGLSLATNPKDNSPLA 120

Qy 121 CSPHSHCGSSVYTTGKSRVNSNPRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180  
Db |||  
Qy 121 CSPHSHCGSSVYTTGKSRVNSNPRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180  
Db |||  
Qy 181 VQFLINILKKEFIIGPQIQVGVVQGEDVHVEFHNDYRSKDVVEAAASHIEQRGTET 240  
Db |||  
Qy 181 VQFLINILKKEFIIGPQIQVGVVQGEDVHVEFHNDYRSKDVVEAAASHIEQRGTET 240  
Db |||  
Qy 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSHDSPDLEKVIQOSERNVTRYAVAVL 300  
Db |||  
Qy 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSHDSPDLEKVIQOSERNVTRYAVAVL 300  
Db |||  
Qy 301 GYNRGINPETFLNEIKYTASDPDDKHPNVTDEAALKDIDVALGDRIPSEGTNNET 360  
Db |||  
Qy 301 GYNRGINPETFLNEIKYTASDPDDKHPNVTDEAALKDIDVALGDRIPSEGTNNET 360  
Db |||  
Qy 361 SFGLEMSOTGFSHVVEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPELKN 420  
Db |||  
Qy 361 SFGLEMSOTGFSHVVEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPELKN 420  
Db |||  
Qy 421 HGAYLGYTVTSVSSROGRVYVAGAPFNHTGKVLFTMNNRSLTIHQAMRQQIGSYF 480  
Db |||  
Qy 421 HGAYLGYTVTSVSSROGRVYVAGAPFNHTGKVLFTMNNRSLTIHQAMRQQIGSYF 480  
Db |||  
Qy 481 GSEITSVDIDGQVTDVLLVGAEMYFNEGRERKGVVYELRQNFVYNGTLKDSHSYQNA 540  
Db |||  
Qy 481 GSEITSVDIDGQVTDVLLVGAEMYFNEGRERKGVVYELRQNFVYNGTLKDSHSYQNA 540  
Db |||  
Qy 541 RFGSSIASVRDLAQDSYNDVVGAPLEDNEAGAIYIFHGRGSIILKTPKQITASELATG 600  
Db |||  
Qy 541 RFGSSIASVRDLAQDSYNDVVGAPLEDNEAGAIYIFHGRGSIILKTPKQITASELATG 600  
Db |||  
Qy 601 LQVFGCSIHGLDNLDELGLDGLAVGALGNVILWSRPVQINASLHPEPSKINIFHRDCK 660  
Db |||  
Qy 601 LQVFGCSIHGLDNLDELGLDGLAVGALGNVILWSRPVQINASLHPEPSKINIFHRDCK 660  
Db |||  
Qy 661 RSGEDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYTPRAHLDGSGDRFTNRAV 720  
Db |||  
Qy 661 RSGEDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYTPRAHLDGSGDRFTNRAV 720  
Db |||  
Qy 721 LLSGGQELCRINHFHLDTDADYKPVTFVSVEYSLDHPDHPKMDCKNPTTLRVSVPFWNG 780  
Db |||  
Qy 721 LLSGGQELCRINHFHLDTDADYKPVTFVSVEYSLDHPDHPKMDCKNPTTLRVSVPFWNG 780  
Db |||  
Qy 781 CNDDEHCVPLDLARSDLPFAMEYCORVLKPAQCSAYTLGDFDTVTFIESTRQAV 840  
Db |||  
Qy 781 CNDDEHCVPLDLARSDLPFAMEYCORVLKPAQCSAYTLGDFDTVTFIESTRQAV 840  
Db |||  
Qy 841 EATLENGENAYSTVLNISANLQFASLIQKEDSDGSIIECVNEERLQKQVNSYPPF 900  
Db |||  
Qy 841 EATLENGENAYSTVLNISANLQFASLIQKEDSDGSIIECVNEERLQKQVNSYPPF 900  
Db |||  
Qy 901 RAKAKVAFRLDSRFSKIFLHLEIELAAGSDSNRSDTKEDNVAPLRPHLYEADVLT 960  
Db |||  
Qy 901 RAKAKVAFRLDSRFSKIFLHLEIELAAGSDSNRSDTKEDNVAPLRPHLYEADVLT 960  
Db |||  
Qy 961 RSSLSHYEKLNSLERYDGIOPPFSCIFRIQNLGLFPIHGMKMTIPIATRSNRL 1020  
Db |||  
Qy 961 RSSLSHYEKLNSLERYDGIOPPFSCIFRIQNLGLFPIHGMKMTIPIATRSNRL 1020  
Db |||  
Qy 1021 KLKDFLTDE-ANTSQNTWGNSTYRPTPVEEDLRAPOLNHSNDVVSINCNRLVPOB 1079  
Db |||  
Qy 1021 KLKDFLTDE-ANTSQNTWGNSTYRPTPVEEDLRAPOLNHSNDVVSINCNRLVPOB 1080  
Db |||  
Qy 1080 INFHLGNLWLRSLKALKYKSMKIMVNAALQROFHSPPFIPEEDPSRQIEFISKQEDMQ 1139  
Db |||  
Qy 1081 INFHLGNLWLRSLKALKYKSMKIMVNAALQROFHSPPFIPEEDPSRQIEFISKQEDMQ 1140  
Db |||  
Qy 1140 VPIWIIVSGTIGLGLLILALVLRKLGFPFSARRRRRREGLDTPKYLE 1188  
Db |||  
Qy 1141 VPIWIIVSGTIGLGLLILALVLRKLGFPFSARRRRRREGLDTPKYLE 1189  
Db |||

## RESULT 11

ADE63570

ID ADE63570 standard; protein; 1189 AA.

XX ADE63570;

XX 29-JAN-2004 (first entry)

XX Human Protein Q9UKX5, SEQ ID NO 9514.

XX Human; pain; neuronal tissue; gene therapy;

XX spinal segmental nerve injury; chronic constriction injury; CCI;

XX spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GHEO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 1189 AA;

Query Match 99.4%; Score 6188.5; DB 7; Length 1189;

Best Local Similarity 99.6%; Fred. No. 0;

Matches 1184; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 1 MDLPRGLVAVNALSLNPGFTDTFNMDTRKPRVPGSRPTAFYTYVQOHDISGNKWLVGGA 60

Db 1 MDLPRGLVVAWALSFWPGFTDFTNMDTRKPRVPGSRATFGYTVQOHDISGNKLVVGA 60  
Qy 61 PLETFNGYOKTGVYACPVTHGNTKLANLGRVTLNVSERKONRGLGLSLATNPKDNSFLA 120  
Db 61 PLETFNGYOKTGVYACPVTHGNTKLANLGRVTLNVSERKONRGLGLSLATNPKDNSFLA 120  
Qy 121 CSPLASHECGSSYTTGMCSSRVNSFRFSKTVAPALQRCQTYMDIVIVLDGNSNIPWVE 180  
Db 121 CSPLASHECGSSYTTGMCSSRVNSFRFSKTVAPALQRCQTYMDIVIVLDGNSNIPWVE 180  
Qy 181 VOHFLINILKKPYIGGQIQGVQYGEDVHVEPHLNDYRSKDVVEAASHIEQGGTET 240  
Db 181 VOHFLINILKKPYIGGQIQGVQYGEDVHVEPHLNDYRSKDVVEAASHIEQGGTET 240  
Qy 241 RTAFGIEFARSAFOKGRGKAKVMIVITDGSHSDSPLEKVIQOSRDNVTRVAVL 300  
Db 241 RTAFGIEFARSAFOKGRGKAKVMIVITDGSHSDSPLEKVIQOSRDNVTRVAVL 300  
Qy 301 GYNRGINPETFLEIKYIASDDPKHFNVTDBAALKDVIDALGDRIFSLGNTKNET 360  
Db 301 GYNRGINPETFLEIKYIASDDPKHFNVTDBAALKDVIDALGDRIFSLGNTKNET 360  
Qy 361 SFGLEMSQTFSSHVVDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPEELKN 420  
Db 361 SFGLEMSQTFSSHVVDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPEELKN 420  
Qy 421 HGAYLGTVTVSVSSRQGVYVAGAPRFNETGKVLFTMNNRSLTTHQAVGGQIGSYF 480  
Db 421 HGAYLGTVTVSVSSRQGVYVAGAPRFNETGKVLFTMNNRSLTTHQAVGGQIGSYF 480  
Qy 481 GSEITSVDIDGQVTDVLGAPMVFNEGRGKVVYVELRQNRVYNGTLKDSYQNA 540  
Db 481 GSEITSVDIDGQVTDVLGAPMVFNEGRGKVVYVELRQNRVYNGTLKDSYQNA 540  
Qy 541 RFGSSIASVRDLNQSDYNDVVGAPLEDNHAGALYIFHGFGRSILKTPKQRTIASSELATG 600  
Db 541 RFGSSIASVRDLNQSDYNDVVGAPLEDNHAGALYIFHGFGRSILKTPKQRTIASSELATG 600  
Qy 601 LOYFGCSIHGQDLNEDGLDIAVAGALGNVILWSRPVQINASHLHFPESKINI 660  
Db 601 LOYFGCSIHGQDLNEDGLDIAVAGALGNVILWSRPVQINASHLHFPESKINI 660  
Qy 661 RSGRATCLAAFLCFTPIFLAHPFQTTTIGIRYNATMDERRYPRAHLDEGGDRTNRAV 720  
Db 661 RSGRATCLAAFLCFTPIFLAHPFQTTTIGIRYNATMDERRYPRAHLDEGGDRTNRAV 720  
Qy 721 LLSGGQELCERINFHVLDTADYVKPVTFVSVEYSLDDPDHGPMLDDGWPTTLRVSPFWNG 780  
Db 721 LLSGGQELCERINFHVLDTADYVKPVTFVSVEYSLDDPDHGPMLDDGWPTTLRVSPFWNG 780  
Qy 781 CNEDEHCVPLVDLARSDLPTAMEYCORVLRKPAQDCSAYTILSPDTTVFIISSTRQAV 840  
Db 781 CNEDEHCVPLVDLARSDLPTAMEYCORVLRKPAQDCSAYTILSPDTTVFIISSTRQAV 840  
Qy 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERRLQKQCNVSPFF 900  
Db 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERRLQKQCNVSPFF 900  
Qy 901 RAKAKVAPLDDSEFSKIFLHLELELAAGSDSNERDSTKEDNVAFLPHLYKYEADVLPT 960  
Db 901 RAKAKVAPLDDSEFSKIFLHLELELAAGSDSNERDSTKEDNVAFLPHLYKYEADVLPT 960  
Qy 961 RSSLSHVEKLNLSLERYDGIQPPFSCIFRIQNLGLPFIHGMWMTIPIATRSNRL 1020  
Db 961 RSSLSHVEKLNLSLERYDGIQPPFSCIFRIQNLGLPFIHGMWMTIPIATRSNRL 1020  
Qy 1021 KLKDLFUTDE-ANTSCNINWNSYETPTPVEEDLRAPOLNHSNDSVINCINRLVPOE 1079  
Db 1021 KLKDLFUTDE-ANTSCNINWNSYETPTPVEEDLRAPOLNHSNDSVINCINRLVPOE 1080  
Qy 1080 INFHLLGNLWLSKALKYKSMKIMVNAALQRFHSPFIFREDPSRQIVFEISKQEDWQ 1139

Db 1081 INFHLLGNLWLSKALKYKSMKIMVNAALQRFHSPFIFREDPSRQIVFEISKQEDWQ 1140  
Qy 1140 VPIWIIYVSGSTGGLLALLVLAARLKGFPFESASRRBEPGLDPTPKVLE 1188  
Db 1141 VPIWIIYVSGSTGGLLALLVLAARLKGFPFESASRRBEPGLDPTPKVLE 1189  
RESULT 12  
ID ADE86584  
AC ADE86584 standard; protein; 1189 AA.  
XX ADE86584;  
XX 29-JAN-2004 (first entry)  
XX Novel human secreted protein #7.  
XX human; secreted protein; cancer; liver disorder; hepatitis;  
KW neural disorder; Alzheimer's disease.  
XX Homo sapiens.  
XX US2003129685-A1.  
XX 10-JUL-2003.  
XX 18-APR-2001; 2001US-00836353.  
XX 28-OCT-1998; 98US-0105971P.  
XX 27-OCT-1999; 99WO-US025031.  
XX 19-APR-2000; 2000US-0198407P.  
XX (NIJJ/) NI J.  
XX (YOUN/) YOUNG P B.  
XX (KENN/) KENNY J J.  
XX (OLSE/) OLSEN H S.  
XX (MOOR/) MOORE P A.  
XX (WEIY/) WEI Y.  
XX (GREB/) GREENE J M.  
XX (RUBE/) RUBEN S M.  
XX Ni J, Young PB, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;  
PI Ruben SM;  
XX WPI; 2004-020335/02.  
XX N-PSDB; ADE86566.  
XX New nucleic acid molecule, useful for preparing a medicament for  
preventing, treating or ameliorating a medical condition e.g. cancer,  
PT liver disorders or neural disorders.  
XX Claim 11; SEQ ID NO 35; 380pp; English.  
XX The invention relates to an isolated nucleic acid sequence, or its  
CC allelic variant, a fragment of the cDNA sequence, or its fragment,  
CC domain, epitope or species homologue. The nucleic acid is useful for  
CC preparing a medicament for preventing, treating or ameliorating a medical  
CC condition e.g., cancer, liver disorders such as hepatitis or neural  
CC disorders such as Alzheimer's disease. The present sequence represents  
CC the amino acid sequence of a novel human secreted protein.  
XX Sequence 1189 AA;  
SQ  
Query Match 99.4%; Score 6188.5; DB 8; Length 1189;  
Best Local Similarity 99.6%; Pred.No. 0;  
Matches 1184; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
Qy 1 MDLPRGLVVAWALSFWPGFTDFTNMDTRKPRVPGSRATFGYTVQOHDISGNKLVVGA 60  
Db 1 MDLPRGLVVAWALSFWPGFTDFTNMDTRKPRVPGSRATFGYTVQOHDISGNKLVVGA 60  
Qy 61 PLETFNGYOKTGVYACPVTHGNTKLANLGRVTLNVSERKONRGLGLSLATNPKDNSFLA 120  
Db 61 PLETFNGYOKTGVYACPVTHGNTKLANLGRVTLNVSERKONRGLGLSLATNPKDNSFLA 120  
Qy 121 CSPLASHECGSSYTTGMCSSRVNSFRFSKTVAPALQRCQTYMDIVIVLDGNSNIPWVE 180  
Db 121 CSPLASHECGSSYTTGMCSSRVNSFRFSKTVAPALQRCQTYMDIVIVLDGNSNIPWVE 180  
Qy 181 VOHFLINILKKPYIGGQIQGVQYGEDVHVEPHLNDYRSKDVVEAASHIEQGGTET 240  
Db 181 VOHFLINILKKPYIGGQIQGVQYGEDVHVEPHLNDYRSKDVVEAASHIEQGGTET 240  
Qy 241 RTAFGIEFARSAFOKGRGKAKVMIVITDGSHSDSPLEKVIQOSRDNVTRVAVL 300  
Db 241 RTAFGIEFARSAFOKGRGKAKVMIVITDGSHSDSPLEKVIQOSRDNVTRVAVL 300  
Qy 301 GYNRGINPETFLEIKYIASDDPKHFNVTDBAALKDVIDALGDRIFSLGNTKNET 360  
Db 301 GYNRGINPETFLEIKYIASDDPKHFNVTDBAALKDVIDALGDRIFSLGNTKNET 360  
Qy 361 SFGLEMSQTFSSHVVDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPEELKN 420  
Db 361 SFGLEMSQTFSSHVVDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPEELKN 420  
Qy 421 HGAYLGTVTVSVSSRQGVYVAGAPRFNETGKVLFTMNNRSLTTHQAVGGQIGSYF 480  
Db 421 HGAYLGTVTVSVSSRQGVYVAGAPRFNETGKVLFTMNNRSLTTHQAVGGQIGSYF 480  
Qy 481 GSEITSVDIDGQVTDVLGAPMVFNEGRGKVVYVELRQNRVYNGTLKDSYQNA 540  
Db 481 GSEITSVDIDGQVTDVLGAPMVFNEGRGKVVYVELRQNRVYNGTLKDSYQNA 540  
Qy 541 RFGSSIASVRDLNQSDYNDVVGAPLEDNHAGALYIFHGFGRSILKTPKQRTIASSELATG 600  
Db 541 RFGSSIASVRDLNQSDYNDVVGAPLEDNHAGALYIFHGFGRSILKTPKQRTIASSELATG 600  
Qy 601 LOYFGCSIHGQDLNEDGLDIAVAGALGNVILWSRPVQINASHLHFPESKINI 660  
Db 601 LOYFGCSIHGQDLNEDGLDIAVAGALGNVILWSRPVQINASHLHFPESKINI 660  
Qy 661 RSGRATCLAAFLCFTPIFLAHPFQTTTIGIRYNATMDERRYPRAHLDEGGDRTNRAV 720  
Db 661 RSGRATCLAAFLCFTPIFLAHPFQTTTIGIRYNATMDERRYPRAHLDEGGDRTNRAV 720  
Qy 721 LLSGGQELCERINFHVLDTADYVKPVTFVSVEYSLDDPDHGPMLDDGWPTTLRVSPFWNG 780  
Db 721 LLSGGQELCERINFHVLDTADYVKPVTFVSVEYSLDDPDHGPMLDDGWPTTLRVSPFWNG 780  
Qy 781 CNEDEHCVPLVDLARSDLPTAMEYCORVLRKPAQDCSAYTILSPDTTVFIISSTRQAV 840  
Db 781 CNEDEHCVPLVDLARSDLPTAMEYCORVLRKPAQDCSAYTILSPDTTVFIISSTRQAV 840  
Qy 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERRLQKQCNVSPFF 900  
Db 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERRLQKQCNVSPFF 900  
Qy 901 RAKAKVAPLDDSEFSKIFLHLELELAAGSDSNERDSTKEDNVAFLPHLYKYEADVLPT 960  
Db 901 RAKAKVAPLDDSEFSKIFLHLELELAAGSDSNERDSTKEDNVAFLPHLYKYEADVLPT 960  
Qy 961 RSSLSHVEKLNLSLERYDGIQPPFSCIFRIQNLGLPFIHGMWMTIPIATRSNRL 1020  
Db 961 RSSLSHVEKLNLSLERYDGIQPPFSCIFRIQNLGLPFIHGMWMTIPIATRSNRL 1020  
Qy 1021 KLKDLFUTDE-ANTSCNINWNSYETPTPVEEDLRAPOLNHSNDSVINCINRLVPOE 1079  
Db 1021 KLKDLFUTDE-ANTSCNINWNSYETPTPVEEDLRAPOLNHSNDSVINCINRLVPOE 1080  
Qy 1080 INFHLLGNLWLSKALKYKSMKIMVNAALQRFHSPFIFREDPSRQIVFEISKQEDWQ 1139

Db 61 PLETNGYQKTDVYKCPVHGNCTKLNGLRVTLSNVSEKDNMRLGLSLATNPKNDSFLA 120  
Qy 121 CSPLWSECCSSYYTTCMCGRVNSNFRFSKTVAPALQRCOTYMDIVIVLDGNSIYPWVE 180  
Db 121 CSPLWSECCSSYYTTCMCGRVNSNFRFSKTVAPALQRCOTYMDIVIVLDGNSIYPWVE 180  
Qy 181 VOHFLINILKXFIYGPQIQGVVQYGEDVHVFHFLNDYRSVDVVEAASHIEQRGSTET 240  
Db 181 VOHFLINILKXFIYGPQIQGVVQYGEDVHVFHFLNDYRSVDVVEAASHIEQRGSTET 240  
Qy 241 RTAFGLFAPSEAFQKGRGAKKMWIVITDGHSDSPDLEKVIQOSERDNVTRYAVAVL 300  
Db 241 RTAFGLFAPSEAFQKGRGAKKMWIVITDGHSDSPDLEKVIQOSERDNVTRYAVAVL 300  
Qy 301 GYNNRGINPETELNEIKYIASDPDDKHFFNVDEAALKDIDVALGDRIFSLGTTKNET 360  
Db 301 GYNNRGINPETELNEIKYIASDPDDKHFFNVDEAALKDIDVALGDRIFSLGTTKNET 360  
Qy 361 SPGLEMSQTFSSHVVDGVLGNVGYDNGAVLAKETSAGKVIPIRESVLKEFPRESLKN 420  
Db 361 SPGLEMSQTFSSHVVDGVLGNVGYDNGAVLAKETSAGKVIPIRESVLKEFPRESLKN 420  
Qy 421 HGAYLGYTVTVSVSSRCGRVYVAGAPRFNHTGKVIPLTFMNNRSLTIHQAMRQQIGSYF 480  
Db 421 HGAYLGYTVTVSVSSRCGRVYVAGAPRFNHTGKVIPLTFMNNRSLTIHQAMRQQIGSYF 480  
Qy 481 GSEITSVDIDGCTVDVLLVGCAPMYFNEGRERKGVVYELRQNFYVNGTLKDSHXYNA 540  
Db 481 GSEITSVDIDGCTVDVLLVGCAPMYFNEGRERKGVVYELRQNFYVNGTLKDSHXYNA 540  
Qy 541 RFGSSIASVRDLNQDSYNDVVVGCAPLEDNHAGAIYIFHGFPGSILKTPKQITASELATG 600  
Db 541 RFGSSIASVRDLNQDSYNDVVVGCAPLEDNHAGAIYIFHGFPGSILKTPKQITASELATG 600  
Qy 601 LQYFGCSIHQCLDNLNDEGLDLAVGALGNVAVILWSPRVQINASLHFEPSKINIFHRDCK 660  
Db 601 LQYFGCSIHQCLDNLNDEGLDLAVGALGNVAVILWSPRVQINASLHFEPSKINIFHRDCK 660  
Qy 661 RSGRDATCLAAFLCFPIFLAPHFQTTVGIWIRNATMDERRYTPRAHLDGGRDFTNRAV 720  
Db 661 RSGRDATCLAAFLCFPIFLAPHFQTTVGIWIRNATMDERRYTPRAHLDGGRDFTNRAV 720  
Qy 721 LLSGGQELCBRIHFVLDADSLTAMEXCORVLRKPAQCSAVTSLSFDTTVFLIESTRORAV 780  
Db 721 LLSGGQELCBRIHFVLDADSLTAMEXCORVLRKPAQCSAVTSLSFDTTVFLIESTRORAV 780  
Qy 781 CNEDEHCVPLVDARSDDLTAPEYCORVLRKPAQCSAVTSLSFDTTVFLIESTRORAV 840  
Db 781 CNEDEHCVPLVDARSDDLTAPEYCORVLRKPAQCSAVTSLSFDTTVFLIESTRORAV 840  
Qy 841 EATLENRGENAYSTVLNLSQSANLQFASLIQKEDSDGSEICVNEERLQKQVCNVSYPFF 900  
Db 841 EATLENRGENAYSTVLNLSQSANLQFASLIQKEDSDGSEICVNEERLQKQVCNVSYPFF 900  
Qy 901 RAKAKVAFRLDSEFSKSIPLHLEIHLAAGSDNERSDKEDNVAPLHLYKYEADVLFT 960  
Db 901 RAKAKVAFRLDSEFSKSIPLHLEIHLAAGSDNERSDKEDNVAPLHLYKYEADVLFT 960  
Qy 961 RSSLSHYEVLKNSLERYDGIQPPFSCIPRIQNLGIFPHGMKMTIPIATRSQNRLL 1020  
Db 961 RSSLSHYEVLKNSLERYDGIQPPFSCIPRIQNLGIFPHGMKMTIPIATRSQNRLL 1020  
Qy 1021 KLKDFLTDE-ANTSCINWNGSTYRTPVEEDLRRAPOLNHSNDSVSNCRNLRLVPOE 1079  
Db 1021 KLKDFLTDEANTSCINWNGSTYRTPVEEDLRRAPOLNHSNDSVSNCRNLRLVPOE 1080  
Qy 1080 INFHLLGNLWLSLKALKYKSMIMVNAALQROFHSPFFREDDPSRQIEFISKQEDWQ 1139  
Db 1081 INFHLLGNLWLSLKALKYKSMIMVNAALQROFHSPFFREDDPSRQIEFISKQEDWQ 1140  
Qy 1140 VPIWIIWVSTGLGILLALLVLAALRKLGFRSARRRREPGLDPTPKVLE 1188  
Db 1141 VPIWIIWVSTGLGILLALLVLAALRKLGFRSARRRREPGLDPTPKVLE 1189

RESULT 13

ABR58365  
ID ABR58365 standard; protein; 1120 AA.

XX

AC ABR58365;

XX DT 07-JUL-2003 (first entry)

XX Human NOV2b.

XX Human; NOV; antidiabetic; anorectic; antibacterial; virucide;

XX immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia;

XX anti-parkinsonian; antilipaeamic; gene therapy; metabolic disorder;

XX diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;

XX neurodegenerative disorder; Alzheimer's disease; immune disorder;

XX haematopoietic disorder.

XX OS Homo sapiens.

XX PN WO2003029423-A2.

XX PD 10-APR-2003.

XX PF 02-OCT-2002; 2002WO-US031358.

XX PR 02-OCT-2001; 2001US-0326483P.

XX PR 05-OCT-2001; 2001US-0327342P.

XX PR 09-OCT-2001; 2001US-0327917P.

XX PR 09-OCT-2001; 2001US-0328029P.

XX PR 09-OCT-2001; 2001US-0328044P.

XX PR 09-OCT-2001; 2001US-0328056P.

XX PR 12-OCT-2001; 2001US-0328849P.

XX PR 15-OCT-2001; 2001US-0329414P.

XX PR 17-OCT-2001; 2001US-0330142P.

XX PR 22-OCT-2001; 2001US-0341058P.

XX PR 24-OCT-2001; 2001US-0339266P.

XX PR 24-OCT-2001; 2001US-0343629P.

XX PR 29-OCT-2001; 2001US-0349575P.

XX PR 01-NOV-2001; 2001US-0346357P.

XX PR 12-APR-2002; 2002US-0371972P.

XX PR 12-APR-2002; 2002US-0371980P.

XX PR 17-APR-2002; 2002US-0373261P.

XX PR 19-APR-2002; 2002US-0373805P.

XX PR 23-APR-2002; 2002US-0374738P.

XX PR 16-MAY-2002; 2002US-0381101P.

XX PR 17-MAY-2002; 2002US-0381635P.

XX PR 29-MAY-2002; 2002US-0383830P.

XX PR 01-OCT-2002; 2002US-00262839.

XX (CURA-) CURAGEN CORP.

XX Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;

XX Edinger SR, Ellerman K, Gerlach VL, Gorman L, Gao X, Ji W;

XX Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;

XX Rothenberg ME, Shinkets RA, Smithson G, Spytek KA, Taupier RJ;

XX Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;

XX WPI; 2003-381625/36.

XX DR N-PSDB; ACC72077.

XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or

XX treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or

XX dyslipidemia, and in chromosome mapping, tissue typing or

XX pharmacogenomics.

XX Claim 1; Page 107; 487pp; English.

XX The present invention relates to novel human NOV proteins and their

XX coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV

XX proteins are useful in manufacturing a medicament for treating a syndrome

XX associated with a human disease. The NOV proteins and coding sequences



CC may be used to diagnose, treat or prevent metabolic disorders such as  
CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative  
CC disorders such as Alzheimer's disease or Parkinson's disease, immune  
CC disorders, haematopoietic disorders and various dyslipidaemias  
XX  
SQ Sequence 1120 AA;

Query Match 92.9%; Score 5780; DB 6; Length 1120;  
Best Local Similarity 93.8%; Pred. NO. 0;  
Matches 1115; Conservative 1; Mismatches 3; Indels 70; Gaps 2;

QY 1 MDLPRGLVAVALSLLWPGFTDTFNDTRKPRVIGSRRTAFPGYTVQOQHDISGNKWLAVGA 60  
DB 1 MDLPRGLVAVALSLLWPGFTDTFNDTRKPRVIGSRRTAFPGYTVQOQHDISGNKWLAVGA 60  
QY 61 PLENGYKQTDGVYKCPVIHGNCTKLNIGRVTLNVSVERKDNMRGLGSLATNPKDNSFLA 120  
DB 61 PLENGYKQTDGVYKCPVIHGNCTKLNIG- 89  
QY 121 CSPLWSHSCGSSYYTTCGSRVNSFRSKTVAPALQRCOTYMDIVIVLDGNSNTPMVE 180  
DB 90 -----COTYMDIVIVLDGNSNTPMVE 111  
QY 181 VQHFLLINILKKFYIGGQIQGVVQYGEDVWHEFLNDYRSVKOVVEAASHIEQGGTET 240  
DB 112 VQHFLLINILKKFYIGGQIQGVVQYGEDVWHEFLNDYRSVKOVVEAASHIEQGGTET 171  
QY 241 RTAFGIEFARSEAFQKGRGKAKKVMIVITGESHDSPLLEKVIQOQSRDNVTRYAVAVL 300  
DB 172 RTAFGIEFARSEAFQKGRGKAKKVMIVITGESHDSPLLEKVIQOQSRDNVTRYAVAVL 231  
QY 301 GYNNRGINPTFNEIKYIASDPDDKHFFNVYDEAALKDIDVALGDRIFSLGNTKNKET 360  
DB 232 GYNNRGINPTFNEIKYIASDPDDKHFFNVYDEAALKDIDVALGDRIFSLGNTKNKET 291  
QY 361 SFGLEMSQTGSSSHVVEBVGVLGAVGAYDMWGANVLKETSAGKVIPLRESYLKEPPELUN 420  
DB 292 SFGLEMSQTGSSSHVVEBVGVLGAVGAYDMWGANVLKETSAGKVIPLRESYLKEPPELUN 351  
QY 421 HGAYLGTVTVSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSLTTHQMRGQOIGSYF 480  
DB 352 HGAYLGTVTVSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSLTTHQMRGQOIGSYF 411  
QY 481 GSEITSDIDGQVTDVLLGAPMVFNEGRGRGVYVYELRQNRFPVNGTLKDSHESYONA 540  
DB 412 GSEITSDIDGQVTDVLLGAPMVFNEGRGRGVYVYELRQNRFPVNGTLKDSHESYONA 471  
QY 541 RFGSSIASVRDLNODSYNDVVVGAFLPDNTHAGATYIFHGFRGSILKTPKQRTASELATG 600  
DB 472 RFGSSIASVRDLNODSYNDVVVGAFLPDNTHAGATYIFHGFRGSILKTPKQRTASELATG 531  
QY 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILASRPVYVQINASLHFPZSKINIFHRDCK 660  
DB 532 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILASRPVYVQINASLHFPZSKINIFHRDCK 591  
QY 661 RSGRDATCLAAFLCPTPIFLAPHFQTTVIGIRYNATMDERYTTPRAHLDEGGDFTRNAV 720  
DB 592 RSGRDATCLAAFLCPTPIFLAPHFQTTVIGIRYNATMDERYTTPRAHLDEGGDFTRNAV 651  
QY 721 LLSQOELCERINPFLVLOTADYVKVPTFSVEYSLEDDPHGPMDDGWPTTLRVSVFPWNG 780  
DB 652 LLSQOELCERINPFLVLOTADYVKVPTFSVEYSLEDDPHGPMDDGWPTTLRVSVFPWNG 711  
QY 781 CNEDEHCVPLVDLARSDLPTAMEYCORVLRKPAQDCSAYTSLSEDTTTFVFIESTQRVAV 840  
DB 712 CNEDEHCVPLVDLARSDLPTAMEYCORVLRKPAQDCSAYTSLSEDTTTFVFIESTQRVAV 771  
QY 841 EATLNRGENAVSTVLNTSOSANLQFASLIOKESDGSIECVNEERLQKQCNVSVFPF 900  
DB 772 EATLNRGENAVSTVLNTSOSANLQFASLIOKESDGSIECVNEERLQKQCNVSVFPF 831  
QY 901 RAKAVAFRLDSFESKSIFLHLEIELAAGSDNSNERDSTKEDNVAPLRFHLKYEADVLFT 960

Db 832 RAKAVAFRLDSFESKSIFLHLEIELAAGSDNSNERDSTKEDNVAPLRFHLKYEADVLFT 891  
QY 961 RSSLSLHYEYKLNSSLERYDGIQPPFSCIFRIQNLGLPPIHGMKMTIPIATRSNRL 1020  
Db 892 RSSLSLHYEYKLNSSLERYDGIQPPFSCIFRIQNLGLPPIHGMKMTIPIATRSNRL 951  
QY 1021 KLRFELTDE-ANTSCNMGNSTEYRPPVEEDLRRAPQLNHSNDDVVSINCNIELVPOE 1079  
Db 952 KLRFELTDEVANTSCNMGNSTEYRPPVEEDLRRAPQLNHSNDDVVSINCNIELVPOE 1011  
QY 1080 INPHLLGNLWLSLKALKYKSMKIMVNAALQRFHSPFIIFREEDPSPQIIFEISKQBDWQ 1139  
Db 1012 INPHLLGNLWLSLKALKYKSMKIMVNAALQRFHSPFIIFREEDPSPQIIFEISKQBDWQ 1071  
QY 1140 VPIWIIIVGSTLGGILLALLVLAIRKLGFPFSARRRREPGLDTPPKYLE 1188  
Db 1072 VPIWIIIVGSTLGGILLALLVLAIRKLGFPFSARRRREPGLDTPPKYLE 1120  
RESULT 14  
AAB50087  
ID AAB50087 standard; protein; 1188 AA.  
XX  
XX AAB50087;  
XX  
DT 19-MAR-2001 (first entry)  
XX  
DE Murine A259.  
XX  
KW Human; A259; integrin alpha subunit; fibrosis; liver disease; lung;  
KW kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;  
KW rheumatoid arthritis.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT Domain 1. .1141  
FT Peptide 1. .22 /label= Extracellular\_domain  
FT Protein 23. .1188 /label= Signal\_peptide  
FT Domain 39. .74 /label= Mature\_protein  
FT Domain /label= Integrin\_alpha\_subunit\_repeat\_domain\_#1  
FT Domain 115. .157 /label= Integrin\_alpha\_subunit\_repeat\_domain\_#2  
FT Domain 164. .345 /label= I domain  
FT Domain 367. .392 /label= Integrin\_alpha\_subunit\_repeat\_domain\_#3  
FT Domain 421. .455 /label= Integrin\_alpha\_subunit\_repeat\_domain\_#4  
FT Domain 478. .516 /label= Integrin\_alpha\_subunit\_repeat\_domain\_#5  
FT Domain 540. .575 /label= Integrin\_alpha\_subunit\_repeat\_domain\_#6  
FT Domain 602. .640 /label= Integrin\_alpha\_subunit\_repeat\_domain\_#7  
FT Domain 1142. .1164 /label= Transmembrane\_domain  
FT Domain 1165. .1188 /label= Cytoplasmic\_domain  
XX  
FN WO200073339-A1.  
XX  
XX 07-DEC-2000.  
XX  
XX 15-MAY-2000; 2000WO-US013262.  
XX  
XX 28-MAY-1999; 99US-00322790.  
XX 27-APR-2000; 2000US-00561263.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
PA



XX	Pan Y, Lora JM;	
XX	WPI; 2001-041142/05.	
DR	N-PSDB; AAC91904, AAC91905.	
XX	Nucleic acid encoding alpha-integrin subunits, useful for treatment and diagnosis of fibrosis, e.g. of the liver.	
PT	Claim 8; Fig 5; 164pp; English.	
XX	The present sequence is murine integrin alpha subunit, A259 is homologous with the alpha1 and alpha10 integrin subunits and is overexpressed in fibrosis. A259 is implicated in regulation of proliferation, differentiation and/or function of many different cell types. Inhibitors of A259 activity are useful for the treatment of liver disease, particularly fibrosis, and also fibrosis in other organs (specifically lung and kidney). In addition, A259 can be used for treatment and prevention of cancer, osteoporosis, acute myeloid leukaemia, HIV infection, and rheumatoid arthritis	
XX	Sequence 1188 AA;	
SQ	Query Match 91.5%; Score 5693; DB 4; Length 1188;	
	Best Local Similarity 90.2%; Pred. No. 0;	
	Matches 1071; Conservative 57; Mismatches 60; Indels 0; Gaps 0;	
QY	1 MDLPRGLVAVWALSMPGFTDTFNMTRKPRVTPGSRRTAFPGYTVQOQHDISGNKMLVWGA 60	
DB	1 MDPFRGLVAVWALSMPGFTDTFNMTRKPRVTPGSRRTAFPGYTVQOQHDISGNKMLVWGA 60	
QY	61 PLBTNGVQKTDYKCPVHGNCTKLNGLRVTLNSVSRKDNRLGLSLATNPKNLSFLA 120	
DB	61 PMETNGHQTKDYKCPVHGNCTKLNGLRVTLNSVSRKDNRLGLSLATNPKNLSFLA 120	
QY	121 CSPLWSEHCSSYTTGMSRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSIYPWYE 180	
DB	121 CSPLWSEHCSSYTTGMSRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSIYPWYE 180	
QY	181 VOHFLNLIKFYIGPQIQVGVVQYGEDVWHFHLNDRSVKDVVAAASHIEQRGTTET 240	
DB	181 VOHFLNLIKFYIGPQIQVGVVQYGEDVWHFHLNDRSVKDVVAAASHIEQRGTTET 240	
QY	241 RTAFGIEFARSEAFQGRGAKKVMIVTDGSHSDPLEKVIQOQSDNDVTRVAVL 300	
DB	241 RTAFGIEFARSEAFQGRGAKKVMIVTDGSHSDPLEKVIQOQSDNDVTRVAVL 300	
QY	301 GYNNRGINPETFLNEIKYIASDDPKHFFNVTDAAALDXIDVALGDRIFPSLEGTKNET 360	
DB	301 GYNNRGINPETFLNEIKYIASDDPKHFFNVTDAAALDXIDVALGDRIFPSLEGTKNET 360	
QY	361 SPGLEMSQTFSSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPPELKN 420	
DB	361 SPGLEMSQTFSSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPPELKN 420	
QY	421 HGAYLGYTVTSVSSRQGRVYVAGAPRNHTGKVIPLTMNNRSLTIHQMRQOQIGSYF 480	
DB	421 HAAYLGYTVTSVSSRQGRVYVAGAPRNHTGKVIPLTMNNRSLTIHQMRQOQIGSYF 480	
QY	481 GSITSDIDGDTVDVLLGAPWYFSEGRGKVVYVYELRQKRFVYVNGTLKDSHSVQNA 540	
DB	481 GSITSDIDGDTVDVLLGAPWYFSEGRGKVVYVYELRQKRFVYVNGTLKDSHSVQNA 540	
QY	541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNAGAIYIFPHGFRGSILKTPKQRTITASELATG 600	
DB	541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNAGAIYIFPHGFRGSILKTPKQRTITASELATG 600	
QY	601 LOYFGCSIHGOLDLNEGLDLDLAVGALGNVILMSRVVQVAINASLHFEPSKINIHRDCK 660	
DB	601 LOYFGCSIHGOLDLNEGLDLDLAVGALGNVILMSRVVQVAINASLHFEPSKINIHRDCK 660	
QY	661 RSGRDATCLAAFLCFTPIFLAPHEQTATVGRYVATMDERRYPRAHLDEGDDFTNRAV 720	

DB	661 RNGRDATCLAAFLCFTPIFLAPHEQTATVGRYVATMDERRYPRAHLDEGDDFTNRAV 720	
QY	721 LLSGQELCERINFHVLDTADYVKPVTFVSVEYSLEDDPHGEMLDGWPFTTLRVSPVPWNG 780	
DB	721 LLSGQELCERINFHVLDTADYVKPVTFVSVEYSLEDDPHGEMLDGWPFTTLRVSPVPWNG 780	
QY	781 CNEDEHCVDPDVLVDARSDLPTAMEYCORVLRKPAQDCSAYTSLSPDTTVPFIIESTRQVAV 840	
DB	781 CNEDEHCVDPDVLVDARSDLPTAMEYCORVLRKPAQDCSAYTSLSPDTTVPFIIESTRQVAV 840	
QY	841 EATLENGENAYSTVLANISQANLQFASLIQKEDSDGSEICVNBERRKQKOVNVSYPFF 900	
DB	841 EATLENGENAYSTVLANISQANLQFASLIQKEDSDGSEICVNBERRKQKOVNVSYPFF 900	
QY	901 RAKAVAFRLDSEFSKSIPLHLHLEIAAGSDNERNSTKEDNVAPLHFLKYEADVLT 960	
DB	901 RAKAVAFRLDSEFSKSIPLHLHLEIAAGSDNERNSTKEDNVAPLHFLKYEADVLT 960	
QY	961 RSSLSHVEVKNLSLEYDGGIPPPSCIFRQNLGLPFIHGMWKIPIPIATSGNELL 1020	
DB	961 RSSLSHVEVKNLSLEYDGGIPPPSCIFRQNLGLPFIHGMWKIPIPIATSGNELL 1020	
QY	1021 KLRDPLTDEANTSCNMGNSTEYRTPVVEEDLRAPQLNHSNSDVVSNCRNLVNPNOEI 1080	
DB	1021 KLRDPLTDEANTSCNMGNSTEYRTPVVEEDLRAPQLNHSNSDVVSNCRNLVNPNOEI 1080	
QY	1081 NPHLLGNLWLSLTKALKYKSMKIMVNAALQOHPSPFIREFEDPSRQIEFISKOEDMQV 1140	
DB	1081 NPHLLGNLWLSLTKALKYKSMKIMVNAALQOHPSPFIREFEDPSRQIEFISKOEDMQV 1140	
QY	1141 PIWIIIVGSTLGLLHLLALLVLAALRKLGFPSRRRRRPEGLDPTPKVLE 1188	
DB	1141 PIWIIIVGSTLGLLHLLALLVLAALRKLGFPSRRRRRPEGLDPTPKVLE 1188	
RESULT 15		
AAU10552		
ID	AAU10552 standard; protein; 1188 AA.	
XX	AAU10552;	
AC	AAU10552;	
DT	14-FEB-2002 (first entry)	
XX	Murine A259 polypeptide.	
DE	Mouse; A259; integrin alpha subunit; integrin alpha 10; secreted protein; liver disease; fibrosis; lung; kidney; bone associated disorder; blood; cartilage associated disorder; haematopoietic disorder; bone marrow; immune related disease; apoptotic disorder; neuronal tissue disease; neurodegenerative disease; gene therapy; cancer; cytostatic; osteopathic; nephrotropic; immunomodulator; anti-inflammatory; neuroprotective; antiarthritic; antianemic; antiallergic; antiasthmatic; dermatological; antidiabetic; anticonvulsant; antiparkinsonian.	
XX	Mus musculus.	
XX	Key	Location/Qualifiers
FT	Domain	1..1141
FT	Peptide	/note= "Extracellular domain"
FT	Protein	1..22
FT	Domain	/note= "Signal peptide"
FT	Domain	23..1188
FT	Domain	/note= "Mature murine A259"
FT	Domain	39..74
FT	Domain	/note= "Integrin alpha repeat domain"
FT	Domain	115..157
FT	Domain	/note= "Integrin alpha repeat domain"
FT	Domain	164..345
FT	Domain	/note= "I domain or Von Willebrand Factor type A domain"
FT	Domain	367..392
FT	Domain	/note= "Integrin alpha repeat domain"
FT	Domain	421..455
FT	Domain	/note= "Integrin alpha repeat domain"

FT Domain 478..516  
 FT /note= "Integrin alpha repeat domain"  
 FT Domain 540..575  
 FT /note= "Integrin alpha repeat domain"  
 FT Domain 602..640  
 FT /note= "Integrin alpha repeat domain"  
 FT Domain 1142..1164  
 FT /note= "Transmembrane domain"  
 FT Domain 1165..1188  
 FT /note= "Cytoplasmic domain"  
 XX  
 PN W0200181414-A2.  
 XX  
 XX 01-NOV-2001.  
 PD  
 XX  
 XX 27-APR-2001; 2001WO-US013516.  
 PF  
 XX 27-APR-2000; 2000US-00561263.  
 PR  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA  
 XX Pan Y, Lora J;  
 PI  
 XX WPI; 2002-041397/05.  
 DR N-PSDB; AAS16874.  
 DR  
 XX  
 XX New A259 nucleic acids and polypeptides, which comprise integrin alpha  
 PT subunit, useful for diagnosing, preventing or treating e.g. liver  
 PT disease, kidney or lung fibrosis, cancers, blood disorders or immune  
 PT related diseases.  
 XX  
 PS Claim 9; Fig 5; 168pp; English.  
 XX  
 XX The invention relates to human and murine A259 nucleic acid molecules  
 CC which encode secreted proteins with homology to integrin alpha subunits,  
 CC specifically to integrin alpha 10. The A259 polypeptide and nucleic acid  
 CC are useful for treating liver disease or fibrosis, particularly kidney  
 CC fibrosis or lung fibrosis. The A259 polypeptide and nucleic acid are also  
 CC useful for diagnosing, preventing or treating cartilage and bone  
 CC associated disorders (such as bone cancer, achondroplasia, myeloma,  
 CC fibrous dysplasia, scoliosis, osteoarthritis, osteosarcoma and  
 CC osteoporosis), bone marrow, blood and haematopoietic disorders (such as  
 CC acute myeloid leukaemia, haemophilia, anaemia and thalassemia), immune  
 CC related diseases (such as HIV, viral infections, cancers, T cell  
 CC autoimmune disorders e.g. AIDS, and allergic inflammatory disorders e.g.  
 CC asthma and psoriasis), apoptotic disorders (such as systemic lupus  
 CC erythematosus and insulin-dependent diabetes mellitus), diseases of the  
 CC neuronal tissues (such as epilepsy and muscular dystrophy) and  
 CC neurodegenerative diseases (such as Parkinson's disease and Huntington's  
 CC disease). This sequence represents the murine A259 polypeptide  
 XX  
 SQ Sequence 1188 AA;  
 Query Match 91.5%; Score 5693; DB 5; Length 1188;  
 Best Local Similarity 90.2%; Fred. No. 0;  
 Matches 1071; Conservative 57; Mismatches 60; Indels 0; Gaps 0;  
 QY 1 MDLPRLVVAWALSHPGFTDFNMTRKPRVPGSRFAFGYTVQOHDISGNKWL VVGA 60  
 DB 1 MDLPRLVVAWALSHPGFTDFNMTRKPRVPGSRFAFGYTVQOHDISGNKWL VVGA 60  
 QY 61 PLETNGYKGTGVYKCPVHNGCTKLNLRVTLNSVSRKDNMRGLSLATNPKDNSFLA 120  
 DB 61 PNETNGHQTGVYKCPVQGNCTKLNLRVTLNSVSRKDNMRGLSLATNPKDNSFLA 120  
 QY 121 CSPLWSHSCGSSYYTTCGMSRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSNTPWVE 180  
 DB 121 CSPLWSHSCGSSYYTTCGMSRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSNTPWVE 180  
 QY 181 VOHFLINILKFPYIGPGIYQVGYGSDVHFEHLNDRSVKGVVWEAASHIEORGGTET 240  
 DB 181 VOHFLINILKFPYIGPGIYQVGYGSDVHFEHLNDRSVKGVVWEAASHIEORGGTET 240

Search completed: June 24, 2004, 17:54:32  
 Job time : 71 secs

QY 241 RTAFGIEFARSEAFQKGRGAKVMIVITDGSBHDSPDLKVIQOSERONVTRYAVAVL 300  
 DB 241 RTAFGIEFARSEAFQKGRGAKVMIVITDGSBHDSPDLKVIQOSERONVTRYAVAVL 300  
 QY 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVVDEAALKDIDVALGDRIIFSLEGTNNET 360  
 DB 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVVDEAALKDIDVALGDRIIFSLEGTNNET 360  
 QY 361 SFGLEMSQTGFSSHVVEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLYKBPBBLKN 420  
 DB 361 SFGLEMSQTGFSSHVVEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLYKBPBBLKN 420  
 QY 421 HGAYLYGTVTVSVSSRQGRVYVAGAPRNFHTKGVLLFTMHNNESLTHQAMRGQIGSYF 480  
 DB 421 HAAYLYGTVTVSVSSRQGRVYVAGAPRNFHTKGVLLFTMHNNESLTHQAMRGQIGSYF 480  
 QY 481 GSEITSDIDGDGVDVLLVGAPMYFNEGRERKGVYVYELRQNRFFVYNGTLKDSHYSQNA 540  
 DB 481 GSEITSDVVDNDRVTDVLLVGAPMYFSEGRERKGVYVYELRQNRFFVYNGTLKDSHYSQNA 540  
 QY 541 RFGSSIASVDLNDQSDYNVVGAPLEDNHAGAIYIFHGFGRGSLTKTPKORITASELATG 600  
 DB 541 RFGSCIASVQDLNDQSDYNVVGAPLEDNHRGAIYIFHGFQTNILKPKMQRITASELATG 600  
 QY 601 LQYFGCSIHQDLINEDGLIDLAVGALGNVAVILWSRVPVQINASLHFEPSKINIFHRDCK 660  
 DB 601 LQHFGCSIHQDLINEDGLVDLAVGALGNVAVILWSRVPVQINASLHFEPSKINIFHRDCK 660  
 QY 661 RSGRDATCLAAFLCFTPIFLAPHQTTVGIRYNATMDERRYPTRAILDEGSDFTTRAV 720  
 DB 661 RNRGRDATCLAAFLCFTPIFLAPHQTATVGIYNATMDERRYPTRAILDEGSDFTTRAV 720  
 QY 721 LLSSGQELCRINFHVLDTADYVYKPTVFSVEYSLEDPDHGMPLDDGPTTLRVSVFPWNG 780  
 DB 721 LLSSGQELCRINFHVLDTADYVYKPTVFSVEYSLEDPDHGMPLDDGPTTLRVSVFPWNG 780  
 QY 781 CNEDEHCVPLDLARSDDLPTAMEYCORVLKPAQDCSAYTLSPDTTVTFIESTRQVAV 840  
 DB 781 CNEDEHCVPLDLARSDDLPTAMEYCORVLKPAQDCSAYTLSPDTTVTFIESTRQVAV 840  
 QY 841 EATLENREGENAYSTVLNISOSANLQFASLICKEDSDGSIKVEERLQKQVNVSYPPF 900  
 DB 841 EATLENREGENAYSAVLNISQSENLQFASLICKEDSDGSIKVEERLQKQVNVSYPPF 900  
 QY 901 RAKAKVAFRLDSEFSSKSIPLHLEIEILAAGSDSNERDSTKEDNVAPLRFHLKYEADVLT 960  
 DB 901 RAKAKVAFRLDSEFSSKSIPLHLEIEILAAGSDSNERDSTKEDNVAPLRFHLKYEADVLT 960  
 QY 961 RSSLSHYVKLNSLERYDGIQPPPSCTRIQNLGLFPIHGMKMTIPIATESGNRL 1020  
 DB 961 RSSLSHYVKLNSLERYDGIQPPPSCTRIQNLGLFPIHGMKMTIPIATESGNRL 1020  
 QY 1021 KLDRDLTDEANTSCNIMGNSTERYPTPVEEDLRRAPQLNHSNSDVWSINCINRLVNPQEI 1080  
 DB 1021 MLRDFDFTDQNTSCNIMGNSTERYPTPVEEDLRRAPQLNHSNSDVWSINCINRLVNPQEI 1080  
 QY 1081 NFHLNGLMRLSKALKYKSKMIMVNAALQORPHSPPIFREEDPSRQIEFISKQEDQV 1140  
 DB 1081 SFYLVGNLMTLSLTKALKYRSKITVNAALQORPHSPPIFREEDPSRQIEFISKQEDQV 1140  
 QY 1141 PIMTIIVGSTIGLGLALLALLVLLALRLKLGFFRSARRRRPGLDPTPKVLE 1188  
 DB 1141 PIMTIIVGSTIGLGLALLALLVLLALRLKLGFFRSARRRRPGLDPTPKVLE 1188

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 17:56:16 ; Search time 54 Seconds  
(without alignments)  
6210.900 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 6224

Sequence: 1 MDLPRGLVWALSLWPGFT.....FRSARRRRPGLDTPKYLE 1188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1163542 seqs, 282313645 residues

Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09D\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	6196	99.6	1188	15	US-10-291-265-338
2	6192	99.5	1188	15	US-10-291-265-810
3	6188.5	99.4	1189	10	US-09-984-130-35
4	6188.5	99.4	1189	10	US-09-984-130-35
5	6188.5	99.4	1189	12	US-10-262-839-4
6	5780	92.9	1120	12	US-10-262-839-6
7	5393	86.5	1034	10	US-09-984-130-43
8	5393	86.5	1034	10	US-09-984-130-43
9	3621	58.2	707	9	US-09-764-870-313
10	3621	58.2	707	14	US-10-125-540-313
11	3182	51.1	688	10	US-09-866-050A-624
12	3113	50.0	696	10	US-09-866-050A-501
13	2518	40.5	545	10	US-09-866-050A-500
14	2200	35.3	437	15	US-10-108-260A-3386
15	2166.5	34.8	1179	12	US-09-918-715-250

16	2159	34.7	1151	10	US-09-984-130-103	Sequence 103, App
17	2159	34.7	1151	10	US-09-984-130-103	Sequence 103, App
18	2149.5	34.5	1180	12	US-09-918-715-307	Sequence 307, App
19	1863	29.9	1181	12	US-10-211-462-187	Sequence 187, App
20	1863	29.9	1181	14	US-10-160-354-2	Sequence 2, Appli
21	1863	29.9	1181	15	US-10-295-027-1286	Sequence 1286, Ap
22	1809	29.1	1147	12	US-10-336-603A-42	Sequence 42, Appli
23	1194.5	19.2	1161	9	US-09-350-259-53	Sequence 53, Appli
24	1194.5	19.2	1161	10	US-09-891-943-53	Sequence 53, Appli
25	1184.5	19.0	1161	9	US-09-350-259-55	Sequence 55, Appli
26	1184.5	19.0	1161	10	US-09-891-943-55	Sequence 55, Appli
27	1181.5	19.0	1155	9	US-09-350-259-46	Sequence 46, Appli
28	1181.5	19.0	1155	10	US-09-891-943-46	Sequence 46, Appli
29	1167.5	18.8	1151	9	US-09-350-259-37	Sequence 37, Appli
30	1167.5	18.8	1151	10	US-09-891-943-37	Sequence 37, Appli
31	1146.5	18.4	1170	9	US-09-945-265-2	Sequence 2, Appli
32	1146.5	18.4	1170	12	US-10-261-164-1	Sequence 1, Appli
33	1142.5	18.4	1161	9	US-09-350-259-99	Sequence 99, Appli
34	1142.5	18.4	1161	10	US-09-891-943-99	Sequence 99, Appli
35	1139	18.3	1161	9	US-09-350-259-2	Sequence 2, Appli
36	1139	18.3	1161	10	US-09-891-943-2	Sequence 2, Appli
37	1136	18.3	1223	16	US-10-408-765A-295	Sequence 295, App
38	1132.5	18.2	1163	15	US-10-116-275-204	Sequence 204, App
39	1103.5	17.7	1163	9	US-09-350-259-4	Sequence 4, Appli
40	1103.5	17.7	1163	10	US-09-891-943-4	Sequence 4, Appli
41	1094	17.6	1153	9	US-09-350-259-3	Sequence 3, Appli
42	1094	17.6	1153	10	US-09-302-481A-1	Sequence 1, Appli
43	1094	17.6	1153	10	US-09-891-943-3	Sequence 3, Appli
44	1094	17.6	1153	14	US-10-144-259-30	Sequence 30, Appli
45	1094	17.6	1153	14	US-10-207-655-176	Sequence 176, App

## ALIGNMENTS

### RESULT 1

US-10-291-265-338  
; Sequence 338, Application US/10291265  
; Publication No. US20030232054A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-017 (785)  
; CURRENT APPLICATION NUMBER: US/10/291,265  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 09/491,404  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 03/617,746  
; PRIOR FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 09/631,451  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 09/633,870  
; PRIOR FILING DATE: 2000-09-15  
; NUMBER OF SEQ ID NOS: 944  
; SOFTWARE: fastseq for Windows Version 3.0  
; SEQ ID NO 338  
; LENGTH: 1188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-291-265-338

Query Match 99.6%; Score 6196; DB 15; Length 1188;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1184; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 MDLPRGLVWALSLWPGFTDTFNNDYRKPRVIGSRTAFPGYTVQHQDISGNKWLVVG 60  
Db 1 MDLPRGLVWALSLWPGFTDTFNNDYRKPRVIGSRTAFPGYTVQHQDISGNKWLVVG 60  
Qy 61 PLETNGYQKTGDYVKCPVHGNCTKLNLRVTLSNVSEKKNRMLGLSLATNPKNDSFLA 120  
Db 61 PLETNGYQKTGDYVKCPVHGNCTKLNLRVTLSNVSEKKNRMLGLSLATNPKNDSFLA 120

Qy	121	CSPLWSECGSSYTTTCMC	SVNSNFRPSKTVAPALOR	CQTYMDIVIVLDGNSIYPWVE	180
Db	121	CSPLWSECGSSYTTTCMC	SVNSNFRPSKTVAPALOR	CQTYMDIVIVLDGNSIYPWVE	180
Qy	181	VOHFLINILKKFYIGPQOI	QGVVQYGEDVYHFFHLDY	SVKDVVVEAASHIEQGGTET	240
Db	181	VOHFLINILKKFYIGPQOI	QGVVQYGEDVYHFFHLDY	SVKDVVVEAASHIEQGGTET	240
Qy	241	RTAFGIEFARSEAFQKGR	GKAKVMIVITDGSNHDSPD	LEKVIQOQSRDNTVRYAVVL	300
Db	241	RTAFGIEFARSEAFQKGR	GKAKVMIVITDGSNHDSPD	LEKVIQOQSRDNTVRYAVVL	300
Qy	301	GYNNRGINPETFLNEIKY	IASDDPKHFFNVTDAAAL	KDIDVALGDRIFSLGTTNKNET	360
Db	301	GYNNRGINPETFLNEIKY	IASDDPKHFFNVTDAAAL	KDIDVALGDRIFSLGTTNKNET	360
Qy	361	SFGLWSECGSSYTTTCMC	SVNSNFRPSKTVAPALOR	CQTYMDIVIVLDGNSIYPWVE	420
Db	361	SFGLWSECGSSYTTTCMC	SVNSNFRPSKTVAPALOR	CQTYMDIVIVLDGNSIYPWVE	420
Qy	421	HGAYLGYTVTSVSSRQGR	VYVAGAPRNFHTGKVI	LFTMNNRSLTIHQAMRQOIGSYF	480
Db	421	HGAYLGYTVTSVSSRQGR	VYVAGAPRNFHTGKVI	LFTMNNRSLTIHQAMRQOIGSYF	480
Qy	481	GSEITSDVIDDGDGVTDL	LVGAPMYFNEGRGRGVV	VYVYELRQNFVYNGTLKDSHSYQNA	540
Db	481	GSEITSDVIDDGDGVTDL	LVGAPMYFNEGRGRGVV	VYVYELRQNFVYNGTLKDSHSYQNA	540
Qy	541	RFSSIASVRDLNODSYND	VVVGAPLBDNHAGAIYI	FHGFGRGSLKTPKQRTASELATG	600
Db	541	RFSSIASVRDLNODSYND	VVVGAPLBDNHAGAIYI	FHGFGRGSLKTPKQRTASELATG	600
Qy	601	LQYFGCSIHGQDLNEDGL	IDLAVGALGNVILWSRP	PVQVQINASHIHPESKINIFHRDCK	660
Db	601	LQYFGCSIHGQDLNEDGL	IDLAVGALGNVILWSRP	PVQVQINASHIHPESKINIFHRDCK	660

RESULT 2				
US-10-291-265-810				
; Sequence 810, Application US/10291265				
; Publication No. US20030232054A1				
; GENERAL INFORMATION:				
; APPLICANT: Hyseq, Inc.				
; APPLICANT: Tang et al				
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides				
; FILE REFERENCE: 21272-017 (785)				
; CURRENT APPLICATION NUMBER: US/10/291,265				
; CURRENT FILING DATE: 2000-01-25				
; PRIOR APPLICATION NUMBER: 09/491,404				
; PRIOR FILING DATE: 2000-01-25				
; PRIOR APPLICATION NUMBER: 09/617,746				
; PRIOR FILING DATE: 2000-07-17				
; PRIOR APPLICATION NUMBER: 09/631,451				
; PRIOR FILING DATE: 2000-08-03				
; PRIOR APPLICATION NUMBER: 09/633,870				
; PRIOR FILING DATE: 2000-09-15				
; NUMBER OF SEQ ID NOS: 944				
; SOFTWARE: FastSeq for Windows Version 3.0				
; SEQ ID NO 810				
; LENGTH: 1188				
; TYPE: PRT				
; ORGANISM: Homo sapiens				
US-10-291-265-810				
Query Match				
Best Local Similarity 99.5%; Score 6192; DB 15; Length 1188;				
Matches 1183; Conservative 0; Mismatches 5; Indels 0; Gaps 0;				
Qy	1	MDLPRGLVAVALS	SWPGFTDTFNNMTRKPRV	IPGSRVTAFFGYTVQOHDISGNKMLVYGA 60
Db	1	MDLPRGLVAVALS	SWPGFTDTFNNMTRKPRV	IPGSRVTAFFGYTVQOHDISGNKMLVYGA 60
Qy	61	PLENTGYOKTGDVVKCPV	IHGNCNKLNLGRVTL	SNVSRKDNMRGLSLATNPKNDSFLA 120
Db	61	PLENTGYOKTGDVVKCPV	IHGNCNKLNLGRVTL	SNVSRKDNMRGLSLATNPKNDSFLA 120
Qy	121	CSPLWSECGSSYTTTCMC	SVNSNFRPSKTVAPALOR	CQTYMDIVIVLDGNSIYPWVE 180
Db	121	CSPLWSECGSSYTTTCMC	SVNSNFRPSKTVAPALOR	CQTYMDIVIVLDGNSIYPWVE 180
Qy	181	VOHFLINILKKFYIGPQOI	QGVVQYGEDVYHFFHLDY	SVKDVVVEAASHIEQGGTET 240
Db	181	VOHFLINILKKFYIGPQOI	QGVVQYGEDVYHFFHLDY	SVKDVVVEAASHIEQGGTET 240
Qy	241	RTAFGIEFARSEAFQKGR	GKAKVMIVITDGSNHDSPD	LEKVIQOQSRDNTVRYAVVL 300
Db	241	RTAFGIEFARSEAFQKGR	GKAKVMIVITDGSNHDSPD	LEKVIQOQSRDNTVRYAVVL 300
Qy	301	GYNNRGINPETFLNEIKY	IASDDPKHFFNVTDAAAL	KDIDVALGDRIFSLGTTNKNET 360
Db	301	GYNNRGINPETFLNEIKY	IASDDPKHFFNVTDAAAL	KDIDVALGDRIFSLGTTNKNET 360
Qy	361	SFGLWSECGSSYTTTCMC	SVNSNFRPSKTVAPALOR	CQTYMDIVIVLDGNSIYPWVE 420
Db	361	SFGLWSECGSSYTTTCMC	SVNSNFRPSKTVAPALOR	CQTYMDIVIVLDGNSIYPWVE 420
Qy	421	HGAYLGYTVTSVSSRQGR	VYVAGAPRNFHTGKVI	LFTMNNRSLTIHQAMRQOIGSYF 480
Db	421	HGAYLGYTVTSVSSRQGR	VYVAGAPRNFHTGKVI	LFTMNNRSLTIHQAMRQOIGSYF 480
Qy	481	GSEITSDVIDDGDGVTDL	LVGAPMYFNEGRGRGVV	VYVYELRQNFVYNGTLKDSHSYQNA 540
Db	481	GSEITSDVIDDGDGVTDL	LVGAPMYFNEGRGRGVV	VYVYELRQNFVYNGTLKDSHSYQNA 540
Qy	541	RFSSIASVRDLNODSYND	VVVGAPLBDNHAGAIYI	FHGFGRGSLKTPKQRTASELATG 600
Db	541	RFSSIASVRDLNODSYND	VVVGAPLBDNHAGAIYI	FHGFGRGSLKTPKQRTASELATG 600
Qy	601	LQYFGCSIHGQDLNEDGL	IDLAVGALGNVILWSRP	PVQVQINASHIHPESKINIFHRDCK 660
Db	601	LQYFGCSIHGQDLNEDGL	IDLAVGALGNVILWSRP	PVQVQINASHIHPESKINIFHRDCK 660

```

Db 601 LQYFCSIHGQDLNEDGLDLAVGALGNVILMSRPVQINASHFEPSPKINIFHRDCK 660
Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVIGIRYNATMDERRYPTRAHLDEGDRFTNRAV 720
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVIGIRYNATMDERRYPTRAHLDEGDRFTNRAV 720
Qy 721 LLSGQELCERINFHVLDTADYVVKVTFVSVEYSLEDPDHGMPLDGCWTTILRVSPFPWG 780
Db 721 LLSGQELCERINFHVLDTADYVVKVTFVSVEYSLEDPDHGMPLDGCWTTILRVSPFPWG 780
Qy 781 CNEDEHCVPLDVLARSDLPTAMEYCORVLRKPAQDCSAYTLSPDTTFFIESTRQVAV 840
Db 781 CNEDEHCVPLDVLARSDLPTAMEYCORVLRKPAQDCSAYTLSPDTTFFIESTRQVAV 840
Qy 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIKCNBERRLQKQVCNVSYPFF 900
Db 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIKCNBERRLQKQVCNVSYPFF 900
Qy 901 RAKAKVAFLDSEFSKSIPLHLEIEILAAGSDSNERDSTKEDNVAPLRFHLKYEADVLT 960
Db 901 RAKAKVAFLDSEFSKSIPLHLEIEILAAGSDSNERDSTKEDNVAPLRFHLKYEADVLT 960
Qy 961 RSSLSHYEVKLNSSLERYDGIQPPFSCIFRIQNLGLPFIHGMKMTIPIATRSNRL 1020
Db 961 RSSLSHYEVKLNSSLERYDGIQPPFSCIFRIQNLGLPFIHGMKMTIPIATRSNRL 1020
Qy 1021 KLKRDFTLDE-ANTSCNIGWNSTEYRPTFVEEDLRRAPQLNHSNDVVSINCNIPLVNPQE 1079
Db 1021 KLKRDFTLDE-ANTSCNIGWNSTEYRPTFVEEDLRRAPQLNHSNDVVSINCNIPLVNPQE 1079
Qy 1080 INFHLGNLWLSLKALKYKMKIMVNAALQRFHSPIFREEDPSRQIEFISKQEDWQV 1140
Db 1080 INFHLGNLWLSLKALKYKMKIMVNAALQRFHSPIFREEDPSRQIEFISKQEDWQV 1140
Qy 1141 PIWIIVCSITLGLLALLVIALKLGFRSARRRRRGLDPTPKVLE 1188
Db 1141 PIWIIVCSITLGLLALLVIALKLGFRSARRRRRGLDPTPKVLE 1188

RESULT 3
US-09-984-130-35
; Sequence 35, Application US/09984130
; Publication No. US20030055231A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-35

Query Match 99.4%; Score 6188.5; DB 10; Length 1189;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1184; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
Qy 1 MDLPRGLVAVWALSHPGFTDTFNMNDRKPRVIPGSRATFFGYTVQOHDISGNKWLVVGA 60

```

```

Db 1 MDLPRGLVAVWALSHPGFTDTFNMNDRKPRVIPGSRATFFGYTVQOHDISGNKWLVVGA 60
Qy 61 PLETNGYQKTGDVYKCPVIHGNCTKLNGLRVTLNSVSRKDNMLGLSLATNPKNSFLA 120
Db 61 PLETNGYQKTGDVYKCPVIHGNCTKLNGLRVTLNSVSRKDNMLGLSLATNPKNSFLA 120
Qy 121 CSPLWHSHECCSSYYTTGKCSRVSNSNFRPSKTVAPALQRCOTYMDIVILVLDGNSIYVWVE 180
Db 121 CSPLWHSHECCSSYYTTGKCSRVSNSNFRPSKTVAPALQRCOTYMDIVILVLDGNSIYVWVE 180
Qy 181 VQFLINILAKFYIGPQOIQGVVVQYGEDVVFHFLNDYRSVKDVAASHIEQGGTET 240
Db 181 VQFLINILAKFYIGPQOIQGVVVQYGEDVVFHFLNDYRSVKDVAASHIEQGGTET 240
Qy 241 RTAPGIEFASSEAFQKGRGAKKVMIVITDGHSDSPDLKVIQOOSRDNVRYAVAVL 300
Db 241 RTAPGIEFASSEAFQKGRGAKKVMIVITDGHSDSPDLKVIQOOSRDNVRYAVAVL 300
Qy 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVITDEAALKDIDVALGDRIFSLGNTKNET 360
Db 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVITDEAALKDIDVALGDRIFSLGNTKNET 360
Qy 361 SFGLMSQTFSSHHVBDGVLGAVGAYDNGAVLKTSAGKVIPLRESYIKPPEELKN 420
Db 361 SFGLMSQTFSSHHVBDGVLGAVGAYDNGAVLKTSAGKVIPLRESYIKPPEELKN 420
Qy 421 HGAVLGYTTSVSSRQRYVVGAPRNTGKVIPLTMNNRSLTIHQMRGQOIGSYF 480
Db 421 HGAVLGYTTSVSSRQRYVVGAPRNTGKVIPLTMNNRSLTIHQMRGQOIGSYF 480
Qy 481 GSEITSVDIDGQDVTDLVVGAPMYFNEGRGKVVYVYELRQNFVYNGTLKDSHSYQNA 540
Db 481 GSEITSVDIDGQDVTDLVVGAPMYFNEGRGKVVYVYELRQNFVYNGTLKDSHSYQNA 540
Qy 541 RFGSSIASVEDLNQDSYNDVVGAPLEDHAGAIYIIPHGRGSLTKPKORITASELATG 600
Db 541 RFGSSIASVEDLNQDSYNDVVGAPLEDHAGAIYIIPHGRGSLTKPKORITASELATG 600
Qy 601 LQYFCSIHGQDLNEDGLDLAVGALGNVILMSRPVQINASHFEPSPKINIFHRDCK 660
Db 601 LQYFCSIHGQDLNEDGLDLAVGALGNVILMSRPVQINASHFEPSPKINIFHRDCK 660
Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVIGIRYNATMDERRYPTRAHLDEGDRFTNRAV 720
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVIGIRYNATMDERRYPTRAHLDEGDRFTNRAV 720
Qy 721 LLSGQELCERINFHVLDTADYVVKVTFVSVEYSLEDPDHGMPLDGCWTTILRVSPFPWG 780
Db 721 LLSGQELCERINFHVLDTADYVVKVTFVSVEYSLEDPDHGMPLDGCWTTILRVSPFPWG 780
Qy 781 CNEDEHCVPLDVLARSDLPTAMEYCORVLRKPAQDCSAYTLSPDTTFFIESTRQVAV 840
Db 781 CNEDEHCVPLDVLARSDLPTAMEYCORVLRKPAQDCSAYTLSPDTTFFIESTRQVAV 840
Qy 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIKCNBERRLQKQVCNVSYPFF 900
Db 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSIKCNBERRLQKQVCNVSYPFF 900
Qy 901 RAKAKVAFLDSEFSKSIPLHLEIEILAAGSDSNERDSTKEDNVAPLRFHLKYEADVLT 960
Db 901 RAKAKVAFLDSEFSKSIPLHLEIEILAAGSDSNERDSTKEDNVAPLRFHLKYEADVLT 960
Qy 961 RSSLSHYEVKLNSSLERYDGIQPPFSCIFRIQNLGLPFIHGMKMTIPIATRSNRL 1020
Db 961 RSSLSHYEVKLNSSLERYDGIQPPFSCIFRIQNLGLPFIHGMKMTIPIATRSNRL 1020
Qy 1021 KLKRDFTLDE-ANTSCNIGWNSTEYRPTFVEEDLRRAPQLNHSNDVVSINCNIPLVNPQE 1079
Db 1021 KLKRDFTLDE-ANTSCNIGWNSTEYRPTFVEEDLRRAPQLNHSNDVVSINCNIPLVNPQE 1079
Qy 1080 INFHLGNLWLSLKALKYKMKIMVNAALQRFHSPIFREEDPSRQIEFISKQEDWQV 1139

```

Db	1081	INFHLGNLWLSRLKALKYKSMKIMVNAALQRFHSPFFIFREBDDPSRQIVFISKQEDWQ	1140
Qy	1140	VPIWIIIVGSTLGGILLALLVLAIRKLGFPSRRRRRPGCLDTPKYLE	1188
Db	1141	VPIWIIIVGSTLGGILLALLVLAIRKLGFPSRRRRRPGCLDTPKYLE	1189
RESULT 4.			
US-09-836-353A-35			
; Sequence 35, Application US/09836353A			
; Publication No. US20030129685A1			
; GENERAL INFORMATION:			
; APPLICANT: Ni et al.			
; TITLE OF INVENTION: 12 Human Secreted Proteins			
; CURRENT APPLICATION NUMBER: US/09/836,353A			
; PRIORITY FILING DATE: 2001-04-18			
; PRIOR FILING DATE: 2000-04-19			
; PRIOR APPLICATION NUMBER: PCT/US99/25031			
; PRIOR FILING DATE: 1999-10-27			
; PRIOR APPLICATION NUMBER: 60/105,971			
; NUMBER OF SEQ ID NOS: 147			
; SOFTWARE: Patent In Ver. 2.0			
; SEQ ID NO 35			
; LENGTH: 1189			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-836-353A-35			
Query Match			
Best Local Similarity 99.4%; Score 6188.5; DB 10; Length 1189;			
Matches 1184; Conservative 1; Mismatches 3; Indels 1; Gaps 1;			
Qy	1	MDLPRGLVAVWALSPLPGFTDTFNMDTRKPRVTPGRTAFPGVTVQOCHDISGNKMLVGA	60
Db	1	MDLPRGLVAVWALSPLPGFTDTFNMDTRKPRVTPGRTAFPGVTVQOCHDISGNKMLVGA	60
Qy	61	PLETNGYQKTDVYKCPVHIGNCTKLNLRVLSNVSRKDNRLGSLATNPKNSTFLA	120
Db	61	PLETNGYQKTDVYKCPVHIGNCTKLNLRVLSNVSRKDNRLGSLATNPKNSTFLA	120
Qy	121	CSPWHEGSSVYTTGMSRVNSNFRSKTVAPALQRCOTYMDIVIVLDGNSIYPWVE	180
Db	121	CSPWHEGSSVYTTGMSRVNSNFRSKTVAPALQRCOTYMDIVIVLDGNSIYPWVE	180
Qy	181	VQHFLLNLKFFYIGPQIQGVVQYGEDVWVEHFLNDYRSKDVVEAAASHIEQGGTET	240
Db	181	VQHFLLNLKFFYIGPQIQGVVQYGEDVWVEHFLNDYRSKDVVEAAASHIEQGGTET	240
Qy	241	RTAFGIEFARSAPOKGGKAKKUMIVITDGESHSDPLEKVIQOESRDNVTRAVAVL	300
Db	241	RTAFGIEFARSAPOKGGKAKKUMIVITDGESHSDPLEKVIQOESRDNVTRAVAVL	300
Qy	301	GYNRRGINPETFLNEIKYIASDDPKHFNVTDEAALXDIVDALGDRIFSLEGTNKNET	360
Db	301	GYNRRGINPETFLNEIKYIASDDPKHFNVTDEAALXDIVDALGDRIFSLEGTNKNET	360
Qy	361	SPGLEMSQTSFSEHVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPELKN	420
Db	361	SPGLEMSQTSFSEHVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPELKN	420
Qy	421	HGAYLGVTYSVSSRQGRVYVAGAPRPNHTGKVLFTMHNRSLSLTHOMRGOQIGSYF	480
Db	421	HGAYLGVTYSVSSRQGRVYVAGAPRPNHTGKVLFTMHNRSLSLTHOMRGOQIGSYF	480
Qy	481	GSBITSDIDGQVTDVLLVAGAPWYFNEGRGKQVYVELQNRVYNGTLKDSHSYQNA	540
Db	481	GSBITSDIDGQVTDVLLVAGAPWYFNEGRGKQVYVELQNRVYNGTLKDSHSYQNA	540
Qy	541	RFSSIASVRLNODSNVYVVGAPLEDNHAGAIYIPIHGRGSIILKTPKQRTASELATG	600

RESULT 5  
US-10-262-839-4  
; Sequence 4, Application US/10262839  
; Publication No. US20040038877A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsobrook, John,  
; APPLICANT: Anderson, David W.,  
; APPLICANT: Boldog, Ferenc,  
; APPLICANT: Burgess, Catherine,  
; APPLICANT: Catterton, Blina,  
; APPLICANT: Edinger, Shlomit,  
; APPLICANT: Ellerman, Karen,  
; APPLICANT: Gerlach, Valerie,  
; APPLICANT: Gorman, Linda,  
; APPLICANT: Guo, Xiaojia,  
; APPLICANT: Ji, Weizhen,  
; APPLICANT: Kekoda, Ramesh,  
; APPLICANT: Lesach, Martin,  
; APPLICANT: Li, Li,  
; APPLICANT: Miller, Charles,  
; APPLICANT: Patturajan, Meera,  
; APPLICANT: Reiger, Daniel,  
; APPLICANT: Rothenberg, Mark,  
; APPLICANT: Shinkets, Richard,  
; APPLICANT: Smithson, Glennda,  
; APPLICANT: Spytek, Kimberly,  
; APPLICANT: Taupier, Raymond, jr.,  
; APPLICANT: Vernet, Corine,  
; APPLICANT: Voss, Edward,  
; APPLICANT: Zerhusen, Brian,

APPLICANT: Zhong, Mei  
 TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
 FILE REFERENCE: 21402-462A  
 CURRENT APPLICATION NUMBER: US/10/262,839  
 CURRENT FILING DATE: 2002-10-01  
 PRIOR APPLICATION NUMBER: 60/326,483  
 PRIOR FILING DATE: 2001-10-02  
 PRIOR APPLICATION NUMBER: 60/327,917  
 PRIOR FILING DATE: 2001-10-09  
 PRIOR APPLICATION NUMBER: 60/328,029  
 PRIOR FILING DATE: 2001-10-09  
 PRIOR APPLICATION NUMBER: 60/328,056  
 PRIOR FILING DATE: 2001-10-09  
 PRIOR APPLICATION NUMBER: 60/381,101  
 PRIOR FILING DATE: 2002-05-16  
 PRIOR APPLICATION NUMBER: 60/371,972  
 PRIOR FILING DATE: 2002-04-12  
 PRIOR APPLICATION NUMBER: 60/327,342  
 PRIOR FILING DATE: 2001-10-05  
 PRIOR APPLICATION NUMBER: 60/328,044  
 PRIOR FILING DATE: 2001-10-09  
 PRIOR APPLICATION NUMBER: 60/328,849  
 PRIOR FILING DATE: 2001-10-12  
 PRIOR APPLICATION NUMBER: 60/374,738  
 PRIOR FILING DATE: 2002-04-23  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 367  
 SOFTWARE: CuraseqList version 0.1  
 SEQ ID NO 4  
 LENGTH: 1189  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-262-839-4

Query Match 99.4%; Score 6188.5; DB 12; Length 1189;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 1184; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy	1	MDPLGLVAVALS	180
Db	1	MDPLGLVAVALS	180
Qy	61	PLETNGYQKTDGVYKCPVHGNCTKLNIGRVTLSNVSRKDNMRGLSLATNPKNDSFLA	120
Db	61	PLETNGYQKTDGVYKCPVHGNCTKLNIGRVTLSNVSRKDNMRGLSLATNPKNDSFLA	120
Qy	121	CSPLWSHCGSSYTTGCMRSVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSNTPWVE	180
Db	121	CSPLWSHCGSSYTTGCMRSVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSNTPWVE	180
Qy	181	VQHFLINILKXPIYIGQIQVGVQYGEDVWVHEFHNDYRSYKOVVEAASHIBORGSTET	240
Db	181	VQHFLINILKXPIYIGQIQVGVQYGEDVWVHEFHNDYRSYKOVVEAASHIBORGSTET	240
Qy	241	RTAPGIEFARSAPFQGRKXGAKKVMIVITDGSHSDSPDLEKVIQOSRDNVRYAVAVL	300
Db	241	RTAPGIEFARSAPFQGRKXGAKKVMIVITDGSHSDSPDLEKVIQOSRDNVRYAVAVL	300
Qy	301	GYNNRGINPFTFLNEIKYIASDDPKHFFNVYTDAAALKDIDVALGDRIFSLGKNNKNET	360
Db	301	GYNNRGINPFTFLNEIKYIASDDPKHFFNVYTDAAALKDIDVALGDRIFSLGKNNKNET	360
Qy	361	SPGLEMSQTFSSHVVDGVLGAVDNGAVLKETSAKVIPLESYKPEELKN	420
Db	361	SPGLEMSQTFSSHVVDGVLGAVDNGAVLKETSAKVIPLESYKPEELKN	420
Qy	421	HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTWNNRSLTIHQAMRGQIGSYF	480
Db	421	HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTWNNRSLTIHQAMRGQIGSYF	480
Qy	481	GSEITSDVDGQVTDVLLVGPMPYFNEGRECKVYVYELQNRPNYNGTLKDSHYONA	540
Db	481	GSEITSDVDGQVTDVLLVGPMPYFNEGRECKVYVYELQNRPNYNGTLKDSHYONA	540

Qy	541	RFGSSIASVRLDNDQSDNVVWGAPLEDNAGAIYIFHGRGSIILTKPKQITASELATG	600
Db	541	RFGSSIASVRLDNDQSDNVVWGAPLEDNAGAIYIFHGRGSIILTKPKQITASELATG	600
Qy	601	LOYFGCSIHGOLDLNEOGLIDLAVGALNAVILWSRPVQINASLHFEPPSKINIFHRDCK	660
Db	601	LOYFGCSIHGOLDLNEOGLIDLAVGALNAVILWSRPVQINASLHFEPPSKINIFHRDCK	660
Qy	661	RSGRDATCLAAFLCPTPIFLAPHQTTTIGIRNATWDERRYPRAHLDDGGDRFTTRAV	720
Db	661	RSGRDATCLAAFLCPTPIFLAPHQTTTIGIRNATWDERRYPRAHLDDGGDRFTTRAV	720
Qy	721	LLSSGQELCSRIHFHVLDTADYVKPTFSVEYSLEDHDPGMLDDGGPTTLRVSPVFWNG	780
Db	721	LLSSGQELCSRIHFHVLDTADYVKPTFSVEYSLEDHDPGMLDDGGPTTLRVSPVFWNG	780
Qy	781	CNEDEHCVPLDLDARSDDLPTAMEYCORVLRKPAQDCSAYTSLPDTTTFVTFIESTRQAV	840
Db	781	CNEDEHCVPLDLDARSDDLPTAMEYCORVLRKPAQDCSAYTSLPDTTTFVTFIESTRQAV	840
Qy	841	ZATLENREGENAYSTVLNISOSANLQFASLIQKEDSDGSIIECVNEERLQKQVCNVSYPFF	900
Db	841	ZATLENREGENAYSTVLNISOSANLQFASLIQKEDSDGSIIECVNEERLQKQVCNVSYPFF	900
Qy	901	RAKAKVAFRLDSEFSKSIPLHLEIHLAAGSDSNERDSTKEDNVAPLRFHLKYEADVLT	960
Db	901	RAKAKVAFRLDSEFSKSIPLHLEIHLAAGSDSNERDSTKEDNVAPLRFHLKYEADVLT	960
Qy	961	RSSLSHYEVKLNSSERYDGIIPPSCIFRIQNLGLFPIHGMWMTITIPATSGNRL	1020
Db	961	RSSLSHYEVKLNSSERYDGIIPPSCIFRIQNLGLFPIHGMWMTITIPATSGNRL	1020
Qy	1021	KLRLDPLTDE-ANTSCNINWGNSTYRPTTVEEDERRAPQLNHSNSDVVSINCNIRLVPNQ	1079
Db	1021	KLRLDPLTDE-ANTSCNINWGNSTYRPTTVEEDERRAPQLNHSNSDVVSINCNIRLVPNQ	1079
Qy	1080	INFHLLGNLWLSLKALKYKSMKIMVNAALQROFHSPIFREEDPSQIIEFISKOEDWQ	1139
Db	1081	INFHLLGNLWLSLKALKYKSMKIMVNAALQROFHSPIFREEDPSQIIEFISKOEDWQ	1139
Qy	1140	VPTWIIWGSTLGGILLALMLALRLKGFPRFRRRRREPGLDPTPKVLE	1188
Db	1141	VPTWIIWGSTLGGILLALMLALRLKGFPRFRRRRREPGLDPTPKVLE	1189

RESULT 6  
 US-10-262-839-6  
 Sequence 6, Application US/10262839  
 Publication No. US20040038877A1  
 GENERAL INFORMATION:  
 APPLICANT: Alsobrook, John,  
 APPLICANT: Anderson, David W.,  
 APPLICANT: Boldog, Ferenc,  
 APPLICANT: Burgess, Catherine,  
 APPLICANT: Catterton, Elina,  
 APPLICANT: Edinger, Shlomit,  
 APPLICANT: Ellerman, Karen,  
 APPLICANT: Gerlach, Valerie,  
 APPLICANT: Guo, Xiaojia,  
 APPLICANT: Ji, Weizhen,  
 APPLICANT: Kekuda, Ramesh,  
 APPLICANT: Leach, Martin,  
 APPLICANT: Li, Li, Charles,  
 APPLICANT: Miller, Charles,  
 APPLICANT: Patturajan, Meera,  
 APPLICANT: Reiger, Daniel,  
 APPLICANT: Rothenberg, Mark,  
 APPLICANT: Shimkets, Richard,  
 APPLICANT: Smithson, Glennda,  
 APPLICANT: Spytek, Kimberly,  
 APPLICANT: Taupier, Raymond, jr.,



; APPLICANT: Vernet, Corine,  
 ; APPLICANT: Voss, Edward,  
 ; APPLICANT: Zethusen, Brian,  
 ; APPLICANT: Zhong, Mei  
 ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
 ; FILE REFERENCE: 21402-462A  
 ; CURRENT APPLICATION NUMBER: US/10/262,839  
 ; CURRENT FILING DATE: 2002-10-01  
 ; PRIOR APPLICATION NUMBER: 60/326,483  
 ; PRIOR FILING DATE: 2001-10-02  
 ; PRIOR APPLICATION NUMBER: 60/327,917  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/328,029  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/328,056  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/381,101  
 ; PRIOR FILING DATE: 2002-05-16  
 ; PRIOR APPLICATION NUMBER: 60/371,972  
 ; PRIOR FILING DATE: 2002-04-12  
 ; PRIOR APPLICATION NUMBER: 60/327,342  
 ; PRIOR FILING DATE: 2001-10-05  
 ; PRIOR APPLICATION NUMBER: 60/328,044  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: 60/328,849  
 ; PRIOR FILING DATE: 2001-10-12  
 ; PRIOR APPLICATION NUMBER: 60/374,738  
 ; PRIOR FILING DATE: 2002-04-23  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 367  
 ; SOFTWARE: CuraSequid version 0.1  
 ; SEQ ID NO 6  
 ; LENGTH: 1120  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-262-839-6  
  
 Query Match 92.9%; Score 5780; DB 12; Length 1120;  
 Best Local Similarity 93.8%; Pred. No. 0;  
 Matches 1115; Conservative 1; Mismatches 3; Indels 70; Gaps 2;  
  
 QY 1 MDLPRGLVVAWALSINLPGFTDFTNMDTRKPRVIGSRTAFPGYTVQQHD:SGNKLWVGA 60  
 DB 1 MDLPRGLVVAWALSINLPGFTDFTNMDTRKPRVIGSRTAFPGYTVQQHD:SGNKLWVGA 60  
 QY 61 PLETNGYQKTDGVYKCPVHGNCTKLALGRVTLNVSERKDNMRLGLSLATNPKNDSFLA 120  
 DB 61 PLETNGYQKTDGVYKCPVHGNCTKLNLG----- 89  
 QY 121 CSLPLWSHECGSSYTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180  
 DB 90 -----CQTYMDIVIVLDGNSIYPWVE 111  
 QY 181 VQHFLINLKEFYIGPGQIQGVVQYGEDVHVHEFLNDYRSVKDWEAASHIEQGGTET 240  
 DB 112 VQHFLINLKEFYIGPGQIQGVVQYGEDVHVHEFLNDYRSVKDWEAASHIEQGGTET 171  
 QY 241 RTAFGIEPAREAFQKGRKAKKVMIVITDGHSDSPDLKVKVQQSERDNVTRYAVAVL 300  
 DB 172 RTAFGIEPAREAFQKGRKAKKVMIVITDGHSDSPDLKVKVQQSERDNVTRYAVAVL 231  
 QY 301 GYNRRGINPEFFLNEIKYIASDPDCKHFFNVTDAAALKDVIDALGRDIFSLGNTKNET 360  
 DB 232 GYNRRGINPEFFLNEIKYIASDDEDDKHFNVTDAAALKDVIDALGRDIFSLGNTKNET 291  
 QY 361 SFGLMSGTGFSSHHVEDGVLLGAVGYDMNGAVLKETSAGKVIPLRESYLKGFPEELKN 420  
 DB 292 SFGLMSGTGFSSHHVEDGVLLGAVGYDMNGAVLKETSAGKVIPLRESYLKGFPEELKN 351  
 QY 421 HGAYLGVTYTSVSSRGRVAVAGAPRNFHTGKVLFTMNNRSLTTHQAMRGQIGSYF 480  
 DB 352 HGAYLGVTYTSVSSRGRVAVAGAPRNFHTGKVLFTMNNRSLTTHQAMRGQIGSYF 411

QY 481 GSEITSDVIDGQYTDVLLVGAAPMYFNEGRERKGVYVVELRONRVFYNGTLDKSHSYQNA 540  
 DB 412 GSEITSDVIDGQYTDVLLVGAAPMYFNEGRERKGVYVVELRONRVFYNGTLDKSHSYQNA 471  
 QY 541 RFGSSIASVRDLNQDSYNDVVVGAFLDNDHAGAIYIFHGFGRSILKTPKQIRITASELATG 600  
 DB 472 RFGSSIASVRDLNQDSYNDVVVGAFLDNDHAGAIYIFHGFGRSILKTPKQIRITASELATG 531  
 QY 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILSRPVVQINASHLFFBPSKINIFHRDCK 660  
 DB 532 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILSRPVVQINASHLFFBPSKINIFHRDCK 591  
 QY 661 RSGRDATCLAAFLCFTPIFLAPHFOTTTVGIIRYNATMDERRYPTRAHLDEGGDRTNRAV 720  
 DB 592 RSGRDATCLAAFLCFTPIFLAPHFOTTTVGIIRYNATMDERRYPTRAHLDEGGDRTNRAV 651  
 QY 721 LLSGQELCERINFHVLDTADYVYKPVTFSEVSLDEDDPHGPMDDGWPPTTLRVSVFFWNG 780  
 DB 652 LLSGQELCERINFHVLDTADYVYKPVTFSEVSLDEDDPHGPMDDGWPPTTLRVSVFFWNG 711  
 QY 781 CNEDEHCVPLVLDARSOLPTAMEYCORVLRKPAODCSAYTILSFDTTVEFIIESTQRVAV 840  
 DB 712 CNEDEHCVPLVLDARSOLPTAMEYCORVLRKPAODCSAYTILSFDTTVEFIIESTQRVAV 771  
 QY 841 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDSGSIECVNEERLQKQVNSYPPF 900  
 DB 772 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDSGSIECVNEERLQKQVNSYPPF 831  
 QY 901 RAKAKVAFRLDSEPSKSIPLHLEIELAGSUSNERDSTKEONVAPLRPHLXKYEADVLFT 960  
 DB 832 RAKAKVAFRLDSEPSKSIPLHLEIELAGSUSNERDSTKEONVAPLRPHLXKYEADVLFT 891  
 QY 961 RSSLSHSHVEKLNLSLERYDGIQPPFSCIFRIQNLGLFPIHGMMKTIPIATRSNRL 1020  
 DB 892 RSSLSHSHVEKLNLSLERYDGIQPPFSCIFRIQNLGLFPIHGMMKTIPIATRSNRL 951  
 QY 1021 KLRDLFTDE-ANTSCNINWGSTERYPTPVEEDLRAPQLNHSNDVVSINCMRLVPOE 1079  
 DB 952 KLRDLFTDEVANTSCNINWGSTERYPTPVEEDLRAPQLNHSNDVVSINCMRLVPOE 1011  
 QY 1080 INFHILGNLWLSLAKALKYKSMKIWNVAALQRFSPPIFRBEDPSRQIEFBISSKQEDWQ 1139  
 DB 1012 INFHILGNLWLSLAKALKYKSMKIWNVAALQRFSPPIFRBEDPSRQIEFBISSKQEDWQ 1071  
 QY 1140 VPIWIIIVGTIGLGLLLALLLVALLKLGFRSARRRREPCLDPTPKVLE 1188  
 DB 1072 VPIWIIIVGTIGLGLLLALLLVALLKLGFRSARRRREPCLDPTPKVLE 1120  
  
 RESULT 7  
 ; US-09-984-130-43  
 ; Sequence 43, Application US/09984130  
 ; Publication No. US20030055231A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ni et al.  
 ; TITLE OF INVENTION: 12 Human Secreted Proteins  
 ; FILE REFERENCE: PF489P2  
 ; CURRENT APPLICATION NUMBER: US/09/984,130  
 ; CURRENT FILING DATE: 2001-10-29  
 ; PRIOR APPLICATION NUMBER: 60/243,792  
 ; PRIOR FILING DATE: 2000-10-30  
 ; PRIOR APPLICATION NUMBER: 09/836,353  
 ; PRIOR FILING DATE: 2001-04-18  
 ; PRIOR APPLICATION NUMBER: 60/198,407  
 ; PRIOR FILING DATE: 2000-04-19  
 ; PRIOR APPLICATION NUMBER: PCT/US99/25031  
 ; PRIOR FILING DATE: 1999-10-27  
 ; PRIOR APPLICATION NUMBER: 60/105,971  
 ; PRIOR FILING DATE: 1998-10-28  
 ; NUMBER OF SEQ ID NOS: 149  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 43  
 ; LENGTH: 1034



[illegible]

```

Db 541 RFSSIASVRLDNDVNDVVGAPLEDNAGAIYIFHGRGSIILKTPKQITASELATG 600
Qy 601 LQYFGCSIHGQDLNEDGLDLAVGALGNVILWSRPVQVNASLHFEPSKINIFHRDCK 660
Db 601 LQYFGCSIHGQDLNEDGLDLAVGALGNVILWSRPVQVNASLHFEPSKINIFHRDCK 660
Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGIRYNATMDERYTPRAHLDGGDRFTNRAV 720
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGIRYNATMDERYTPRAHLDGGDRFTNRAV 720
Qy 721 LSSGQELCERINPHVLDATYVYKPVTFSEYSELEDPDHGPMDDGMPPTTLRLVSVPFWNG 780
Db 721 LSSGQELCERINPHVLDATYVYKPVTFSEYSELEDPDHGPMDDGMPPTTLRLVSVPFWNG 780
Qy 781 CNEDEHCVPLVDARSDLPTAMEYCORVLKPAQCCSAVTLSDFTTPTLIESSTORVAV 840
Db 781 CNEDEHCVPLVDARSDLPTAMEYCORVLKPAQCCSAVTLSDFTTPTLIESSTORVAV 840
Qy 841 EATLENKGENAYSTVLNISQSANLQFASLIQKSDSDGSIIECVNEERRLOKQVCNVSYPFF 900
Db 841 EATLENKGENAYSTVLNISQSANLQFASLIQKSDSDGSIIECVNEERRLOKQVCNVSYPFF 900
Qy 901 RAKAKVAPRLDSBESKSI FLHLEIELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFT 960
Db 901 RAKAKVAPRLDSBESKSI FLHLEIELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFT 960
Qy 961 RSSLSHYEVKLANSLERYDGIQPPFSCIERIQLGLFPIHGMMKITTPIATRSNRL 1020
Db 961 RSSLSHYEVKLANSLERYDGIQPPFSCIERIQLGLFPIHGMMKITTPIATRSNRL 1020
Qy 1021 KLKRDFLTDE 1029
Db 1021 KLKRDFLTDE 1029

```

RESULT 9

```

US-09-764-870-313
; Sequence 313, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 313
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-313

```

```

Query Match 58.2%; Score 3621; DB 9; Length 707;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 689; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDLPRLVVAWALSLLWPGFTDTFNMDTRKPRVPGSRTAFPGYTVQOHDISGNKWLWVGA 60
Db 17 MDLPRLVVAWALSLLWPGFTDTFNMDTRKPRVPGSRTAFPGYTVQOHDISGNKWLWVGA 76
Qy 61 PLETNGYOKTGDVYKCPVIHGNCTKLNLRVTLSNVSEKDNKRLGLSLATNPKNSFLA 120
Db 77 PLETNGYOKTGDVYKCPVIHGNCTKLNLRVTLSNVSEKDNKRLGLSLATNPKNSFLA 136
Qy 121 CSPLMSHSCGSSYTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNLYPWVE 180
Db 137 CSPLMSHSCGSSYTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNLYPWVE 196
Qy 181 VQHFLINILKKFYIGPGQIQGVQYGEDVVEHFLNDYRSVKDVVEAASHIEQRGGTET 240

```

```

Db 197 VQHFLINILKKFYIGPGQIQGVQYGEDVVEHFLNDYRSVKDVVEAASHIEQRGGTET 256
Qy 241 RTAFGIEFARSAFQKGRKGAKKVMIVITDGESHSDPLEKVIQQSERDNTVRYAVVL 300
Db 257 RTAFGIEFARSAFQKGRKGAKKVMIVITDGESHSDPLEKVIQQSERDNTVRYAVVL 316
Qy 301 GYNNRRGINPETFLNEIKYIASDPDDKHFFPNVTVDEAALKDIDVADLGDRIFFSLGEGTNKNET 360
Db 317 GYNNRRGINPETFLNEIKYIASDPDDKHFFPNVTVDEAALKDIDVADLGDRIFFSLGEGTNKNET 376
Qy 361 SFGLEMSGTGFSSSHVVEDGVLGAVGYDMNGAVLKETISAGKVIPLRSYLYKEFPEELKN 420
Db 377 SFGLEMSGTGFSSSHVVEDGVLGAVGYDMNGAVLKETISAGKVIPLRSYLYKEFPEELKN 436
Qy 421 HGAYLGYTVTSVVSRRQGRVYVAGAPRNFHTKGVILFTWNNRSITIHQAMGQQOIGSYF 480
Db 437 HGAYLGYTVTSVVSRRQGRVYVAGAPRNFHTKGVILFTWNNRSITIHQAMGQQOIGSYF 496
Qy 481 GSEITSDVIDGVDVTVLLVGAPMYFNEGRERKGVYVYELRQNRVYVYNGTLKDSHSYQNA 540
Db 497 GSEITSDVIDGVDVTVLLVGAPMYFNEGRERKGVYVYELRQNRVYVYNGTLKDSHSYQNA 556
Qy 541 REGSSIASVRDLNQDSYNDVVGAPLEDNAGAIYIFHGRGSIILKTPKQITASELATG 600
Db 557 REGSSIASVRDLNQDSYNDVVGAPLEDNAGAIYIFHGRGSIILKTPKQITASELATG 616
Qy 601 LQYFGCSIHGQDLNEDGLDLAVGALGNVILWSRPVQVNASLHFEPSKINIFHRDCK 660
Db 617 LQYFGCSIHGQDLNEDGLDLAVGALGNVILWSRPVQVNASLHFEPSKINIFHRDCK 676
Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGI 691
Db 677 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGI 707

```

RESULT 10

```

US-10-125-540-313
; Sequence 313, Application US/10125540
; Publication No. US20030059875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214C1
; CURRENT APPLICATION NUMBER: US/10/125,540
; CURRENT FILING DATE: 2002-04-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 313
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-125-540-313

```

```

Query Match 58.2%; Score 3621; DB 14; Length 707;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 689; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDLPRLVVAWALSLLWPGFTDTFNMDTRKPRVPGSRTAFPGYTVQOHDISGNKWLWVGA 60
Db 17 MDLPRLVVAWALSLLWPGFTDTFNMDTRKPRVPGSRTAFPGYTVQOHDISGNKWLWVGA 76
Qy 61 PLETNGYOKTGDVYKCPVIHGNCTKLNLRVTLSNVSEKDNKRLGLSLATNPKNSFLA 120
Db 77 PLETNGYOKTGDVYKCPVIHGNCTKLNLRVTLSNVSEKDNKRLGLSLATNPKNSFLA 136
Qy 121 CSPLMSHSCGSSYTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNLYPWVE 180
Db 137 CSPLMSHSCGSSYTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNLYPWVE 196
Qy 181 VQHFLINILKKFYIGPGQIQGVQYGEDVVEHFLNDYRSVKDVVEAASHIEQRGGTET 240
Db 197 VQHFLINILKKFYIGPGQIQGVQYGEDVVEHFLNDYRSVKDVVEAASHIEQRGGTET 256

```

QY 241 RTAFGIEPARSEAFQGRKGAKKVMIVITDGHSDSPDLBKVIOOSREDNVTYAVAVL 300  
DB 257 RTAFGIEPARSEAFQGRKGAKKVMIVITDGHSDSPDLBKVIOOSREDNVTYAVAVL 316  
QY 301 GYNRRGINPFLNEIKVIASDPDDKHFNFVNTDEAALKDVIDALGDRIPSELTGNET 360  
DB 317 GYNRRGINPFLNEIKVIASDPDDKHFNFVNTDEAALKDVIDALGDRIPSELTGNET 376  
QY 361 SFGLEMSQTFSSHVVYDGVLLGAGYDNGAVLKETSAGKVIPLRSYLKPEPPELKN 420  
DB 377 SFGLEMSQTFSSHVVYDGVLLGAGYDNGAVLKETSAGKVIPLRSYLKPEPPELKN 436  
QY 421 HGAYLGYTTSVSSRQGRVYVAGAPRFNHTGKVLFTMHNNRSLTIHOAMEGOOIGSYF 480  
DB 437 HGAYLGYTTSVSSRQGRVYVAGAPRFNHTGKVLFTMHNNRSLTIHOAMEGOOIGSYF 496  
QY 481 GSEITSDVDGGVTDVLLGAPMYNEGREKGVVYELRQNRFTYNGTLKXDSHYQNA 540  
DB 497 GSEITSDVDGGVTDVLLGAPMYNEGREKGVVYELRQNRFTYNGTLKXDSHYQNA 556  
QY 541 RFGSSIASVRDLNODSYNDVVVVGAPLEDNHAGAIYIFHGFSGSIILKTPKQRTITASELATG 600  
DB 557 RFGSSIASVRDLNODSYNDVVVVGAPLEDNHAGAIYIFHGFSGSIILKTPKQRTITASELATG 616  
QY 601 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILWSPVQVQINASLHFSPSKINI FHRDCK 660  
DB 617 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILWSPVQVQINASLHFSPSKINI FHRDCK 676  
QY 661 RSGRDATCLAAFLCFTPIFLAPHPFTTTVGI 691  
DB 677 RSGRDATCLAAFLCFTPIFLAPHPFTTTVGI 707

## RESULT 11

US-09-866-050A-624

; Sequence 624, Application US/09866050A

; Publication No. US2003004071A1

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Murison, James G.

; APPLICANT: Kumble, Krishanand D.

; TITLE OF INVENTION: Compositions Isolated From Skin Cells

; TITLE OF INVENTION: and Methods for Their Use

; FILE REFERENCE: 11000.1011c4U

; CURRENT FILING DATE: 2001-05-24

; NUMBER OF SEQ ID NOS: 725

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 624

; LENGTH: 688

; TYPE: PRT

; ORGANISM: Rat

US-09-866-050A-624

Query Match 51.1%; Score 3182; DB 10; Length 688;

Best Local Similarity 86.5%; Pred. No. 2.6e-284;

Matches 595; Conservative 46; Mismatches 47; Indels 0; Gaps 0;

QY 501 GAPMYFNEGERGVVYVYELRQNRFTYNGTLKXDSHYQNAFSGSSIASVRDLNODSYNDV 560

DB 1 GAPMYFNEGERGVVYVYELRQNRFTYNGTLKXDSHYQNAFSGSSIASVRDLNODSYNDV 60

QY 561 VVGAPLEDNHAGAIYIFHGFSGSIILKTPKQRTITASELATGLOYPGCSIHGQDLNEDGLI 620

DB 61 VVGAPLEDNHAGAIYIFHGFSGSIILKTPKQRTITASELATGLOYPGCSIHGQDLNEDGLV 120

QY 621 DLAVGALGNVILWSPVQVQINASLHFSPSKINI FHRDCKKSGRDATCLAAFLCFTPIFL 680

DB 121 DLAVGALGNVILWSPVQVQINASLHFSPSKINI FHRDCKKSGRDATCLAAFLCFTPIFL 180

QY 681 APHECTTIVGIRYNATMDREYTTBRAHLDEGGDFTNRAVLLSSGOELCERINFHVLDTA 740  
DB 181 APHECTTIVGIRYNATMDREYTTBRAHLDEGGDFTNRAVLLSSGOELCERINFHVLDTA 240  
QY 741 DYVXPVTFVSYSLEDPDHGMPLDDGMPTTLRVSPVFWNGCNBDEHCHVDPDLVLDARS DLP 800  
DB 241 DYVXPVTFVSYSLEDPDHGMPLDDGMPTTLRVSPVFWNGCNBDEHCHVDPDLVLDARS DLP 300  
QY 801 TAMEYCOVRVLRKPAQDCSAVTLSPDTTVFIIESTRORVAVATLENRCENAYSTVLTISQ 860  
DB 301 TAMEYCOVRVLRKPAQDCSAVTLSPDTTVFIIESTRORVAVATLENRCENAYSTVLTISQ 360  
QY 861 SANLQFASLIQKEDSDGSIIECVNEERLQKQCNVSPVFPFRAKAKVAFRLDSEFSKSIPL 920  
DB 361 SENLQFASLIQKEDSDGSIIECVNEERLQKQCNVSPVFPFRAKAKVAFRLDSEFSKSIPL 420  
QY 921 HHLXIELAAGSDNSHERDSTKEDNVAPLRFHLKYEADVLFTRSSSLSHYEVKLNSSLEPYD 980  
DB 421 HHLXIELAAGSDNSHERDSTKEDNVAPLRFHLKYEADVLFTRSSSLSHYEVKLNSSLEPYD 480  
QY 981 GIGPPFCIFRQNLGLFPPIHGMWKITPIATSGNRLLKLRDPLTDENTSCNMGNS 1040  
DB 481 GIGPPFCIFRQNLGLFPPIHGMWKITPIATSGNRLLKLRDPLTDENTSCNMGNS 540  
QY 1041 TEYRPTPVEDLRAPQNLNSNSDVVSINCNIRLVNQBENHFLGNLMLSLKALKYKS 1100  
DB 541 TEYRPTPVEDLRAPQNLNSNSDVVSINCNIRLVNQBENHFLGNLMLSLKALKYKS 600  
QY 1101 MKIMVNAALQRFHSPRIFREEDPSRQIEFEISKOEDQVPIWIIIVGSLGLLALLV 1160  
DB 601 MKIMVNAALQRFHSPRIFREEDPSRQIEFEISKOEDQVPIWIIIVGSLGLLALLV 660  
QY 1161 LALRKLGFRRSARRRERPGDPTPKVLE 1188  
DB 661 LALRKLGFRRSARRRERPGDPTPKVLE 688

## RESULT 12

US-09-866-050A-501

; Sequence 501, Application US/09866050A

; Publication No. US2003004071A1

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Murison, James G.

; APPLICANT: Kumble, Krishanand D.

; TITLE OF INVENTION: Compositions Isolated From Skin Cells

; TITLE OF INVENTION: and Methods for Their Use

; FILE REFERENCE: 11000.1011c4U

; CURRENT APPLICATION NUMBER: US/09/866,050A

; CURRENT FILING DATE: 2001-05-24

; NUMBER OF SEQ ID NOS: 725

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 501

; LENGTH: 696

; TYPE: PRT

; ORGANISM: Rat

US-09-866-050A-501

Query Match 50.0%; Score 3113; DB 10; Length 696;

Best Local Similarity 85.3%; Pred. No. 6.4e-278;

Matches 584; Conservative 45; Mismatches 56; Indels 0; Gaps 0;

QY 501 GAPMYFNEGERGVVYVYELRQNRFTYNGTLKXDSHYQNAFSGSSIASVRDLNODSYNDV 560

DB 1 GAPMYFNEGERGVVYVYELRQNRFTYNGTLKXDSHYQNAFSGSSIASVRDLNODSYNDV 60

QY 561 VVGAPLEDNHAGAIYIFHGFSGSIILKTPKQRTITASELATGLOYPGCSIHGQDLNEDGLI 620

DB 61 VVGAPLEDNHAGAIYIFHGFSGSIILKTPKQRTITASELATGLOYPGCSIHGQDLNEDGLV 120

QY 621 DLAVGALGNVILWSRVPVOINASLHFEPSKINI FHRDCKRSGEDATCLAAFLCPTPIEL 680  
DB 121 DLAVGALGNVILWSRVPVOINASLHFEPSKINI FHRDCKRSGEDATCLAAFLCPTPIEL 180  
QY 681 APHFQTTVGIRYNATMDERRYTPRAHLDEGGDFTNRAVLSSGQELCERINFHVLDTA 740  
DB 181 APHFQTTVGIRYNATMDERRYTPRAHLDEGGDFTNRAVLSSGQELCERINFHVLDTA 240  
QY 741 DYVKPVTESVEYSLEDPDHGMLDGPPTTLRVSVFPWNGCNEDEKCVPLVLDARSJLP 800  
DB 241 DYVKPVAFSVEYSLEDPDHGMLDGPPTTLRVSVFPWNGCNEDEKCVPLVLDARSJLP 300  
QY 801 TAMEYQVRKPAQCSAYTSLFDTTVPFIIESTRQVAVAEATLENRGENAYSTVLNLSQ 860  
DB 301 TAMEYQVRKPAQCSAYTSLFDTTVPFIIESTRQVAVAEATLENRGENAYSTVLNLSQ 360  
QY 861 SANLQASLIQKEDSDGSECTEVEERLQKQVNCVNSYPPFRKAKAVAFRLDSEFSKSIPL 920  
DB 361 SANLQASLIQKEDSDGSECTEVEERLQKQVNCVNSYPPFRKAKAVAFRLDSEFSKSVFL 420  
QY 921 HHLIEILAAGSDNERDSTYKEDNVAPLRPHLYEADVLFTRSSLSHSHYVKNLSLERYD 980  
DB 421 HHLQIHLGAGSDSHEQSDADNDTALLRPHLYEADVLFTRSSLSHSHYVKNLSLERYD 480  
QY 981 GIGPPSCIPRIQNLGLFPIHGMKITTPIATRSNRLKLRDPLTDEANTSCNIWNS 1040  
DB 481 GIGPPNCFVKQNLGFFPIHGMKITTPIATRSNRLKLRDPLTDEANTSCNIWNS 540  
QY 1041 TEYRTPVEEDLRAPOLHNSDDVYSINCNRLVQNOEINPHLLNMLWSLKLKYS 1100  
DB 541 TEYRTPVEEDLRAPOLHNSDDVYSINCNRLVQNOEINPHLLNMLWSLKLKYS 600  
QY 1101 MKITMNAALQROPHSPFIREDPSRQIPEFSKQEDWQVPIWIVGSTLGLLLALLV 1160  
DB 601 MKITMNAALQROPHSPFIREDPSRQIPEFSKQEDWQVPIWIVGSTLGLLLALLV 660  
QY 1161 LALRKLGFESARRRRERGLDTPK 1185  
DB 661 LALGSLVSLKVPASRGPAPWAPSPK 685

## RESULT 13

US-09-866-050A-500  
; Sequence 500, Application US/09866050A  
; Publication No. US200300471A1

## GENERAL INFORMATION:

; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; TITLE OF INVENTION: and Methods for Their Use  
; FILE REFERENCE: 11000.1011C4U  
; CURRENT APPLICATION NUMBER: US/09/866,050A  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 500  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-866-050A-500

Query Match 40.5%; Score 2518; DB 10; Length 545;  
Best Local Similarity 86.2%; Pred. No. 4, 2e-223;  
Matches 470; Conservative 35; Mismatches 40; Indels 0; Gaps 0;

QY 644 SLHFEPKINIFHRDCKRSGEDATCLAAFLCPTPIELAPHFQTTVGIRYNATMDERRY 703  
DB 1 SLHFEPKINIFHRDCKRSGEDATCLAAFLCPTPIELAPHFQTTVGIRYNATMDERRY 60

QY 704 PRAHLEGGDFTNRAVLSSGQELCERINFHVLDTADYVXKVTESVEYSLEDDPHGML 763  
DB 61 PRAHLEGGDFTNRAVLSSGQELCERINFHVLDTADYVXKVTESVEYSLEDDPHGML 120  
QY 764 DDGWFPTTLRVSVFPWNGCNEDEKCVPLVLDARSJLP TAMEYQVRKPAQCSAYTSL 823  
DB 121 DDGWFPTTLRVSVFPWNGCNEDEKCVPLVLDARSJLP TAMEYQVRKPAQCSAYTSL 180  
QY 824 FDTTVPFIIESTRQVAVAEATLENRGENAYSTVLNLSQ SANLQASLIQKEDSDGSECTE 883  
DB 181 FDTTVPFIIESTRQVAVAEATLENRGENAYSTVLNLSQ SANLQASLIQKEDSDGSECTE 240  
QY 884 EERLQKQVNCVNSYPPFRKAKAVAFRLDSEFSKSIPL HHLIEILAAGSDNERDSTYKEDN 943  
DB 241 EERLQKQVNCVNSYPPFRKAKAVAFRLDSEFSKSIPL HHLIEILAAGSDNERDSTYKEDN 300  
QY 944 VAPLRPHLYEADVLFTRSSLSHSHYVKNLSLERYD GIGPPSCIPRIQNLGLFPIHGM 1003  
DB 301 TALLRPHLYEADVLFTRSSLSHSHYVKNLSLERYD GIGPPSCIPRIQNLGLFPIHGM 360  
QY 1004 MKITPIATRSNRLKLRDPLTDEANTSCNIWNS TEYRTPVEEDLRAPOLHNS 1063  
DB 361 MKITPIATRSNRLKLRDPLTDEANTSCNIWNS TEYRTPVEEDLRAPOLHNS 420  
QY 1064 DVSNINCNRLVQNOEINPHLLNMLWSLKLKYS MKITMNAALQROPHSPFIRED 1123  
DB 421 DVSNINCNRLVQNOEINPHLLNMLWSLKLKYS MKITMNAALQROPHSPFIRED 480  
QY 1124 PSRQIPEFSKQEDWQVPIWIVGSTLGLLLALLV LALRKLGFESARRRRERGLDPT 1183  
DB 481 PSRQIPEFSKQEDWQVPIWIVGSTLGLLLALLV LALRKLGFESARRRRERGLDPT 540  
QY 1184 PKVLE 1188  
DB 541 PKVLE 545

## RESULT 14

US-10-108-260A-3386  
; Sequence 3386, Application US/10108260A  
; Publication No. US20040005560A1

## GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1el full length CDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 3386  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-3386

Query Match 35.3%; Score 2200; DB 15; Length 437;  
Best Local Similarity 99.5%; Pred. No. 7e-194;  
Matches 417; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDLPRGLVAVWALSINPFGTDTFNMDTRKPRVPGSRTAFPGYTVQOHDISGNKLVVGA 60  
DB 1 MDLPRGLVAVWALSINPFGTDTFNMDTRKPRVPGSRTAFPGYTVQOHDISGNKLVVGA 60  
QY 61 PLETNQYQKTGDVYKCPVIHGNCTKLNLRVTLNSVSRKDNMELGLSLATNPKDNGSFLA 120  
DB 61 PLETNQYQKTGDVYKCPVIHGNCTKLNLRVTLNSVSRKDNMELGLSLATNPKDNGSFLA 120  
QY 121 CSPWASHCEGSSYTTTCMCRVNSNFRFSKTVPALORCQTTMDIVIVLQSGNSIYPWVE 180  
DB 121 CSPWASHCEGSSYTTTCMCRVNSNFRFSKTVPALORCQTTMDIVIVLQSGNSIYPWVE 180  
QY 181 VQFLINILKFKYIGPQOIQGVVVOYGEDVVVHPEHLNDYRSVKDWEAAASHI EQRGGTET 240

Db 181 VQHFLNLIKXYIGPQIQVGVVQYGEDVWHEFLNDYRSVKDVEAASHIEQGGTET 240  
Qy 241 RTAFGIEFARSAFQGRGKAKKVMIVITDGHSDSPLEKVIQOOSERDNVTRVAVVL 300  
Db 241 RTAFGIEFARSAFQGRGKAKKVMIVITDGHSDSPLEKVIQOOSERDNVTRVAVVL 300  
Qy 301 GYNNRRGINPFTFLNEIKYIASDPDDKHFFNVNDEAALKDIIDALGDRIFSLEGTKNET 360  
Db 301 GYNNRRGINPFTFLNEIKYIASDPDDKHFFNVNDEAALKDIIDALGDRIFSLEGTKNET 360  
Qy 361 SGLKMSQTFSSHVVEDVILGAVGAYDNGAVLKETSAKVIPIRESYLKEPPELK 419  
Db 361 SFGLEMSQTFSSHVVEDVILGAVGAYDNGAVLKETSAKVIPIRESYLKEPPELK 419

## RESULT 15

US-09-918-715-250  
; Sequence 250, Application US/09918715  
; Publication No. US20030017157A1  
; GENERAL INFORMATION:  
; APPLICANT: Brad St. Croix  
; APPLICANT: Bert Vogelstein  
; APPLICANT: Kenneth Kinzler  
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS  
; FILE REFERENCE: 1107.00134  
; CURRENT APPLICATION NUMBER: US/09/918,715  
; CURRENT FILING DATE: 2001-08-01  
; PRIOR APPLICATION NUMBER: 60/222,599  
; PRIOR FILING DATE: 2000-08-02  
; PRIOR APPLICATION NUMBER: 60/224,360  
; PRIOR FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: 60/282,850  
; PRIOR FILING DATE: 2000-04-11  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 250  
; LENGTH: 1179  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-918-715-250

Query Match 34.8%; Score 2166.5; DB 12; Length 1179;  
Best Local Similarity 38.1%; Pred. No. 4.9e-190;  
Matches 464; Conservative 239; Mismatches 425; Indels 89; Gaps 22;

Qy 4 PR---GLVVA--WALSPLFGFTDTFMNTRKPRVPGSRTAPPFGYTVQOHDISGNKVLV 58  
Db 5 PRARPGVAVACMLLTVVLRCVSPNVDYKNSMTFSGPVEDMFGYTVQOYNEBKWVLI 64  
Qy 59 GAPLETNGYQKTDGVYKCPVIHGN---CTKMLG-RVTLNSVSEKDNKRLGLSLATNPK 114  
Db 65 GSPLVQOPKRTGTYKCPVGRGESLPCVKLDLPVNTSIPNTEVKENKTFGSTLVNTP- 123  
Qy 115 DNSFLACSLPMSHEGSSYTTGMSRVNSNFRPSKTVPALQRCQTMIVILVDGNS 174  
Db 124 NGGFLACGLYAYRCGLHYTTGICSDVSPTQVNVSNAP-VQECSTQLDIVILVDGNS 182  
Qy 175 IYVWVEVQHFLNLIKXYIGPQIQVGVVQYGEDVWHEFLNDYRSVKDVEAASHIEQ 234  
Db 183 IYVWVSVAFLNDLKRMDIGPKQTVQVIGENVNTHFNFLNKYSSTEELVAAKKIVQ 242  
Qy 235 RGGTTRTAFGIEFARSAF--QKGRGAKKVMIVITDGHSDSPLEKVIQOOSERDNV 292  
Db 243 RGGRTMTALGTDTKARKEAFTARGARRGVKKVMIVITDGHSDHNRLLKVIQDCEDENI 302  
Qy 293 TVYAVAVLGYNNRRGINPETFLNEIKYIASDPDDKHFFNVNDEAALKDIIDALGDRIFSL 352  
Db 303 QKPSFALGSGYNNRRGNSLSTKFEVIEKSTASEPETEKHPFNVSDELALVTIVTKLGERIFAL 362  
Qy 353 EGT--NKNETSFGLEMSQTFSSHVVEDVILGAVGAYDNGAVLKETSAKVIPIRESYL 411  
Db 363 EATADQSAASFEMENXQTFGSAHYSDQWMLGAVDNGVTVVWQKASQIIIPRNTTFEN 422

Search completed: June 24, 2004, 18:02:34  
Job time : 58 secs

Qy 412 KEFPEELKNHGAUYLYGTYTTSVVSQRQGRVYVAGAPRNMHTGKVLFTMNNRSLSIHQAM 471  
Db 423 VESTKNEPLASYLYGTYVNSATASSGDVLYIAGQPRNMHTGQVLIYRMDG-NIKILQTL 481  
Qy 472 RQOIGSYFGSEITTSVIDGCVTDVLLVGLAPMTF-NEGRERGKVVVYVELQONFVYNGT 530  
Db 482 SGEQIGSYFGSILTTTIDKDSNTDILLVGLAPMTYMGTEKEBQGVVYVALNOTFPEYQMS 541  
Qy 531 LK-----DSHSYQN-----ARFGSSIASVREDLNQDSYNDVNVVUGAPLEDNEHAG 572  
Db 542 LEPIKQTCSSRQNSCTTENKNEPCGARGTAIAAKDNLNLDGFDNIVIGAPLEDHGG 601  
Qy 573 AIYTFHGRSILTKPKQRTASBELATGLQVFGSISGQLDNLNDELGLDLAVGALGNVAVI 632  
Db 602 AVYIYHSGSKTIRKEYAQRIPSGGDGKTLKPFQGSIRGEMDNLNGDGLTDTVTIGLGGAL 661  
Qy 633 LMSRPVVQINASHLHFBPSKINIHRDCKRSORDATCLAAFLCFTPIFLAPHFQTTVGIR 692  
Db 662 FWRDVAVKVVMFENKVNIOKKNCHMEGKFTVCINATVCFEVKLKSKEDTIEADLQ 721  
Qy 693 YNATMDERRYTPRAHLDGEGDRFTNRVAVLSSGGSELGERINHVLDIADYVYKPTFVSVEY 752  
Db 722 YRVTLDSLRLQISRSFFSGTOERKVQRMITVRKSE--CTKHSFYMLDKHDFQDSVRITLDF 779  
Qy 753 SLEDPDHGPMDDCGPMTTLRVSVFPWNGKNEDEHCVDELVDLARSDLPTAMEYCORLKR 812  
Db 780 NLTDPENGPVLDDSLPNSVHEYIIPFAXDCGKCKISDL----- 818  
Qy 813 PAQDCSAYTISFDTTVEIESTRQRAVEATLENRGENAYSTVLNISQSANLQFASL--I 870  
Db 819 -----SLHVATTEKDLIVRSQNDKFNVSILTAKNKDSAYNRTIVHVSPLVFGIEAI 873  
Qy 871 QKEDSDSIECVNEERLQKQVCNVSYPFPAKAKVAFRLDSEFSKSFILHHLIELAAG 930  
Db 874 QKDSCESN-----HNITCKVGPFLRRGEMVTFKILFQNTSYLKENVTIYLSAT 923  
Qy 931 SDSNERDSTKEDNVAPLRFHLKYEADVLFTRSSLSHVEVKLNS-----LERYDGIQGPFF 986  
Db 924 SDSEPPETLSDNVVNISIPVKYEVGLQFYSSASEYHISIAANETVPEVINSTEDIGNEI 983  
Qy 987 SCITRIONLGLPPIHGMKMITPIATRSNRLKLRDLFTDEANTSC-----NIMGNS 1040  
Db 984 NIFYLRKSGSPFPMPELKLISFPNFTSNGYPVL-YPTGLSSSENANCRPHIFEDPFSIN 1042  
Qy 1041 TEYRPTVEEDLREAPOLAHNSDVSINCNIRLVPNQEINFHLGLMLRLSLKALKYKS 1100  
Db 1043 SKKQTTSTDLKKGTLDCNTCKPATITCMLTSSDISQVNSLI--LWKPTPIKSYFSS 1100  
Qy 1101 MKIMVNAALQRFHSPFIFREEDPSRQTEFISKQE-DWQVPIHVIIVGSTLGGLLALL 1159  
Db 1101 LNLITIRGEL-RSENASLVSSNQKRELAQISKDGLPGRVPLWILLISAPAGLLMLL 1159  
Qy 1160 VLARKLGFERSARRR 1176  
Db 1160 ILALWKIGPFRPLKKK 1176

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 17:53:06 ; Search time 23 Seconds  
(without alignments)

Title: US-09-980-403-2  
 Perfect score: 6224  
 Sequence: 1 MDLPRGLVVAWALSLWPFFT.....FSAARRRREFGLDPTPKVLE 1188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

```
Database : Issued_Patents_AA.*
1: /cgm2_6/pdata/2/iaa/5A_COMB pep.*
2: /cgm2_6/pdata/2/iaa/5B_COMB pep.*
3: /cgm2_6/pdata/2/iaa/6A_COMB pep.*
4: /cgm2_6/pdata/2/iaa/6B_COMB pep.*
5: /cgm2_6/pdata/2/iaa/PTUTS_COMB pep.*
6: /cgm2_6/pdata/2/iaa/backfiles pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2046	32.9	1183	4	US-09-532-3103-5	Sequence 5, Appli
2	1813.5	29.1	1183	4	US-09-532-3103-6	Sequence 6, Appli
3	1194.5	19.2	1161	1	US-08-485-618-53	Sequence 53, Appl
4	1194.5	19.2	1161	1	US-08-362-652-53	Sequence 53, Appl
5	1194.5	19.2	1161	2	US-08-605-672-53	Sequence 53, Appl
6	1194.5	19.2	1161	2	US-08-482-293A-53	Sequence 53, Appl
7	1194.5	19.2	1161	2	US-08-943-363-53	Sequence 53, Appl
8	1194.5	19.2	1161	3	US-09-193-043-53	Sequence 53, Appl
9	1194.5	19.2	1161	4	US-09-688-307A-53	Sequence 53, Appl
10	1194.5	19.2	1161	4	US-09-350-259-53	Sequence 53, Appl
11	1184.5	19.0	1161	3	US-09-193-043-55	Sequence 55, Appl
12	1184.5	19.0	1161	4	US-09-688-307A-55	Sequence 55, Appl
13	1184.5	19.0	1161	4	US-09-350-259-55	Sequence 55, Appl
14	1181.5	19.0	1155	1	US-08-286-889-46	Sequence 46, Appl
15	1181.5	19.0	1155	1	US-08-485-618-46	Sequence 46, Appl
16	1181.5	19.0	1155	1	US-08-362-652-46	Sequence 46, Appl
17	1181.5	19.0	1155	2	US-08-605-672-46	Sequence 46, Appl
18	1181.5	19.0	1155	2	US-08-482-293A-46	Sequence 46, Appl
19	1181.5	19.0	1155	2	US-08-943-363-46	Sequence 46, Appl
20	1181.5	19.0	1155	3	US-09-193-043-46	Sequence 46, Appl
21	1181.5	19.0	1155	4	US-09-688-307A-46	Sequence 46, Appl
22	1181.5	19.0	1155	4	US-09-350-259-46	Sequence 46, Appl
23	1178.5	18.9	1161	1	US-08-485-618-55	Sequence 55, Appl
24	1178.5	18.9	1161	1	US-08-362-652-55	Sequence 55, Appl
25	1178.5	18.9	1161	2	US-08-605-672-55	Sequence 55, Appl
26	1178.5	18.9	1161	2	US-08-482-293A-55	Sequence 55, Appl
27	1178.5	18.9	1161	2	US-08-943-363-55	Sequence 55, Appl

28	1167.5	18.8	1151	1	US-08-286-889-37	Sequence 37, Appl
29	1167.5	18.8	1151	1	US-08-485-618-37	Sequence 37, Appl
30	1167.5	18.8	1151	1	US-08-362-652-37	Sequence 37, Appl
31	1167.5	18.8	1151	1	US-08-605-672-37	Sequence 37, Appl
32	1167.5	18.8	1151	2	US-08-482-293A-37	Sequence 37, Appl
33	1167.5	18.8	1151	2	US-08-482-363-37	Sequence 37, Appl
34	1167.5	18.8	1151	3	US-09-193-043-37	Sequence 37, Appl
35	1167.5	18.8	1151	4	US-09-688-307A-37	Sequence 37, Appl
36	1167.5	18.8	1151	4	US-09-350-259-37	Sequence 37, Appl
37	1146.5	18.4	1170	2	US-08-789-078-2	Sequence 2, Appl
38	1146.5	18.4	1170	2	US-08-752-633-2	Sequence 2, Appl
39	1146.5	18.4	1170	5	PCT-US95-04886-2	Sequence 2, Appl
40	1142.5	18.4	1161	1	US-08-485-618-99	Sequence 99, Appl
41	1142.5	18.4	1161	2	US-08-605-672-99	Sequence 99, Appl
42	1142.5	18.4	1161	2	US-08-482-293A-99	Sequence 99, Appl
43	1142.5	18.4	1161	2	US-08-943-363-99	Sequence 99, Appl
44	1142.5	18.4	1161	3	US-09-193-043-99	Sequence 99, Appl
45	1142.5	18.4	1161	4	US-09-688-307A-99	Sequence 99, Appl

## ALIGNMENTS

RESULT 1  
US-09-532-310B-5  
; Sequence 5, Application US/09532310B  
; Patent No. 6596276  
; GENERAL INFORMATION:  
; APPLICANT: Senger, Donald R  
; Detmar, Michael  
; Claffey, Kevin P  
; TITLE OF INVENTION: Method for inhibiting tumor  
; angiogenesis in a living subject  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David Prashker, Esq.

ZIP: 01930  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.40 Mb storage

COMPUTER: DELL PC  
OPERATING SYSTEM: MS DOS  
SOFTWARE: Microsoft Word version 87

SOFTWARE: MICROSOFT WORD VERSION 9  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: IIS/09/532 310B

APPLICATION NUMBER: 00/00,000,000  
 FILING DATE: 22-Mar-2000  
 CLASSIFICATION: Unknown

ATTORNEY/AGENT INFORMATION:  
NAME: David Prashker, Esq.

REGISTRATION NUMBER: 29,693  
REFERENCE/DOCKET NUMBER: BIS-036

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (978) 525-3794

INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1183 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
-310B-5

atch  
cal Similarity  
32.9%; Score 2046; DB 4; Length 1183;  
36.9%; Pred. No. 2.2e-175;

446; Conservative 233; Mismatches 441; Indels 88; Gaps 21;

11 WALSLWPGFTDFNMDTRKPRVPGSRTAFFGYTVQOHDISGNKWLTVGAPLENTNGYQKT 70

19 WLLTVILGFCVSFNVDVKNSMTFSGPVEDMFGYTVQQYENEEGKKWLIGSPLVGQPKNRT 78



Db 248 DLTTFGAIQARKYASAAAGRRSATKVVVVVTDGSHDGLKAVIDQCNHNTLRF 307  
Qy 236 AVAVLYNNRRGINPE-TFLNEIKYIASDPDDKFFNFVTDAAALKDVIDALGDRIFSLG 354  
Db 308 GIAVLGYNLRNADTKMLIKSEIKALASIPTEYFFNFVSDAALLKXAGTLGEQIFSI 367  
Qy 355 TNKNETSGLEMSQTGFSSHV--VEDGVLLGAVAYDNGAVLKEISAGKVIPLRESYLK 412  
Db 368 TVXGDNFXMMSQVGFSDYSSQNDIILMLGAVGAFGWSGTIVQKTSKHGLI----- 419  
Qy 413 EFPBEL-----KNHGAYLGYTVTSWSSROCRVAVAGAPRFNHTGKVLFTMHNRSLT 466  
Db 420 -FPQAQEOIOIQRNESSVLGVV-AAISTGSTHFVAGAPRANTGQIVLYSVNENGIT 477  
Qy 467 HQAMGGOQISYGFSEITSDVDGVDVLLVCAPMYFNE-GRERGVVYVYBLRNRF 525  
Db 478 VIOHRGQOIGSYGFSLVCSVDVDKDTITDVLVCAPMYSDLKKEGKRVVLFITKIGIL 537  
Qy 526 VYNGTLKXSHYQVAFSSIASVVDLNDQSDVNDVVCAPLEDNAGAIYIFHFGRSIL 585  
Db 538 QHOFLEGEPIENTRFGSARALSDINDGDFNDVIVGSPLENGNSGAVIYINGHGTIR 597  
Qy 596 KIPKQIRITASBLA--TGQYFGCSIHGQLDNEGLDLAVGALGNVILMSRPVQINA 643  
Db 598 TKYSQKILGSDGAFRSHLQYFGSLDGYDLNGDSITDVSIGAFQGVVQLMSQSIAV 657  
Qy 644 SLRPEPSKINIFHRDCKSGRDATCLAFICFTPIELAPHQTTVTGIRYNATWD----E 699  
Db 658 EASTPEKITILVNNAQ-----IILKCFSAKP-RPTQNNQVAVIYNTILDADGFS 708  
Qy 700 RRYTPRAHLDEGDRFTNRVALLSGGELCERINFHVLDTADYVVKPV--TFSVEYSLDDPD 758  
Db 709 SRVTSRGLFKENNERCLQKNVNVQAQSCPEHI-IYIQEPSDVNSLDDLRLVDLSLENG 767  
Qy 759 HGPMLDDGWPTLTVSVFPGWNCDEHCVDPDLV--DARSLPTAMEYQORVLRKPAQC 817  
Db 768 TSPALAYSETAKVFSIFPHXDCGEDGKISDLVLQDVR--IPAAQE-----QP---- 814  
Qy 818 SAYTLSPDTTFTIESTQRVAVEATLENRGENAYSTVLNISQGANLQFASLIQEDSDG 877  
Db 815 -----FIVSNQNKRLTFSTVLNKRKESAYNTGIIVDFSENFFASFLPVD--- 860  
Qy 878 SIECVNEERRLOQV--CNVSYPPFRKAKVAFRLDSBFSKSIPLHLEIELAAGSDSNER 936  
Db 861 GTEVTCQVAASQKSVACDVGYPALKREQQVTFINFDENLQNLQNAQSLSPQALSSEOE 920  
Qy 937 DSTKEDVAPLRFHLKYEADVLFTSSSLSHYEV----KANSRLERYDGIQPPFSCIPRI 992  
Db 921 N--KADNLVNLIKIFLYDAEHLFRTSTNINFTISSGNGVPSIVHSEFEDVGPKF--IFSL 976  
Qy 993 Q-NLGLFPIHGMKMITIPIATRSNRLKLRLDFLTDREANTSNCINWGNSTEYRTPV--- 1048  
Db 977 KVTGSGVPSMATVIIHIPOYTKENPLMYLTGVQTDKAGDIC---CNADINPLKIGOT 1032  
Qy 1049 -----EEDLRAPQLNHSNDVVSINCNRLVNPQINPHILG-----NLWLSL 1093  
Db 1033 SSVSVFSKENFRHTKELNCRATSCSNVTCWLK-----DYHMKGEYFVNVTTRWNGTF 1085  
Qy 1094 KALKYKSKIMVNAALQFSPFFREDEFSRIEPEISK-QEDMOWPIWIIIVGSLIGG 1152  
Db 1086 ASSIFQT--VQLTAAAEINTYNPFIYVIEDNTVPIPLMKPDEKAEVPTGVIGSLIAG 1143  
Qy 1153 LLLALLVALRKLGFFR 1170  
Db 1144 ILLALLVALMLKLGFFK 1161

## RESULT 3

US-08-485-618-53  
; Sequence 53, Application US/08485618  
; Patent No. 5728533  
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Garstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,618  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: US 08/173,497  
APPLICATION NUMBER: 23-DEC-1993  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32797  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 53:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1161 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-485-618-53

Query Match 19.2%; Score 1194.5; DB 1; Length 1161;  
Best Local Similarity 30.1%; Pred. No. 2.3e-98;  
Matches 374; Conservative 198; Mismatches 507; Indels 165; Gaps 49;

Qy 5 RGLVW---AWALSLWFGTDTFMMDTRKRVIPGSRITAPGTYVQQHDISGNKWLAVGAP 61  
Db 3 RGVVILCGWALASCHG----SNLDVEKP-VVFKDAASGQTVVQP---GGSLVVGAP 54  
Qy 62 LETNGYQKTGDVYKCPVHGNCTKLNIGRVTLNVSRKDNMLGLSLATNPKNDSFLAC 121  
Db 55 LEAVAVNQTSQSDCCPACVCCPILL-HIPLAV-----NMSLSLSLVADTNSQLLAC 108  
Qy 122 SPVWSHECGSGSYTTGMCNRVNSNFRPSKTVAPALORCQ-TYNDIVIVLQGSNSI--YFW 178  
Db 109 GPTAQACAKMYKAGSKLLGSSLOFQAIPATMPECPCQEMDIAFLIDGSGSIDQSF 168  
Qy 179 VEVQHFLINILKXFYIGPGOI-----QVGVOVQGEDVHFHNLNDYRSVXDVVEAASHIE 233  
Db 169 TQMKDFVKALM-----GQLASTSTSFSLMQYSLNLIKHTFTTEPKSSLSQSLVDAIV 221  
Qy 234 QRGTEFTRTAFGIEFARSEAP--QKGRKGAKKVMIVITDGBSHDSP-DLEKVIQGBERD 290  
Db 222 QLQGL-TYTAGIQKVVKELFHSKNGARKSAKILIVITDQKPRDPLRYRHIPEAKA 280  
Qy 291 NVTRYAVAVLYNNRRGINPETFLNEIKYIASDPDDKFFNFVTDAAALKDVIDALGDRIF 350  
Db 281 GIIRYAIQVGDAPRE-----PTALQELNTIGSAPSQDHVFKGVNFVALRSIQRIQEKIF 335





Db 449 EVRGTOIGSYFGASLCSDMDRGGSTDLVLIGPHYEHTR--GGQSVSCMPGVRSRWHC 507  
Qy 528 NGTLKDSHSHYQONARFSSIASVRDLNODSYNDVWVVGAPLENEAGAVIFHG--FRGSILK 586  
Db 508 GTTILGEGQHPGRGAAITVLGDVNGDSLADVAIGAPGEEENRGAIVIFHGASQDIAP 567  
Qy 587 TPQRITASELATGQYFQCSHIGLDLNEGLIDLAVGALGNAMVILMSRPVQINASLH 646  
Db 568 SPQRVTGSLFLQYFQCSLGGQDLTQDGLAVLAVGAQGHVLLSLPLKVGISIR 627  
Qy 647 REPSKINIPHRCKSGRDTCL---AAELCT---PIFLAPHFQTTVGIRNATWD 698  
Db 628 FAPSEVAKTVYQC--WGRPTVLEAGATVUTVRKGSDDLIG---DVQSSVRVDLALD 681  
Qy 699 ERYTPRAHLDDGGDFRFTNRAVLLSSGQELCERINFHVLN--TADYVVKPVTFSVEYSL---754  
Db 682 PGRILISRAIFDETCKTLTFRRTGLGDH--CEMKLLLPDCVEDAVTPIILRLNLSLAD 740  
Qy 755 EDPDHG--PMLDDGHTTLRVSPFWNGCNEDEHCVPLDVLDRSDLTAMCYCORVLRK 812  
Db 741 SAPSRNLRLVAVGSDHVTAGFPPEKNCQKELLCEGNL-----779  
Qy 813 PAQDCSAVTLSPDTTTFIETSTRQVAVBEATLENGENAYSTVLNISQSANLOFASLIQK 872  
Db 780 -----GVSNFSGLVLEVGSPELTIVTVWNEGSDSYGLIKFYYPAELSYRVTFA 833  
Qy 873 EDSGD---SIECVNE---ERRLOKQVNVSPFPRAKAVAFRLDSRFSKSIPL--HLEI 925  
Db 834 QQPHYPLRLACEABPTGQESLRSSCSINHPIFREGAKATMITFDVSYKAFLDRLLL 893  
Qy 926 ELAAGSDSNERSTKEDNVAPLPHLYEADVLFRSS-----SLSHVEVLNLSLE 977  
Db 894 RASASENKPISK--TAPQLELPKVTIVTVISROEDSTKHFPSSSHGE--ROKRAEH 950  
Qy 978 RYDGPPSPCIRIQLGLFPIHGMMKXITTIATRSNRL--LKLWDFLTDEANTSCN 1035  
Db 951 RY-----RVNLSPLTL-AISVNFVPIILL-NGVAVMDVTLR---SPAQGVSC- 993  
Qy 1036 IWNSTYRPTPVEDLRAPQ-----LNHSNDVVSINCNI-RVFNQBIHFLLGNL---1088  
Db 994 -----VSOREPPQHSDLLTQIGRSVLDCAIADCHLRCDIPSLGTLDLDELFIKGNLSF 1048  
Qy 1089 -WLRLS---KALKYKSMKIMVNAALQROFSPFIEREEDPSRQIBEFISKQEDMOVPTWI 1144  
Db 1049 GWISLTQKVLLESEAEITNTSVYSQLPQGEAFRAQVSTMLEEYVVB-----PVFL 1103  
Qy 1145 IVGSLTGLLLIALLVLRKLKGFPSARRRE-----PCLDPTP 1184  
Db 1104 MYFSSVGGLLLALITVALYKLGPFK--ROYKEMLDLPSADFP? 1145

RESULT 5

US-08-605-672-53  
; Sequence 53, Application US/08605672  
; Patent No. 5817515  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/605,672  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jz., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 53:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1161 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-605-672-53  
  
Query Match 19.2%; Score 1194.5; DB 2; Length 1161;  
Best Local Similarity 30.1%; Pred. No. 2.3e-98;  
Matches 374; Conservative 198; Mismatches 507; Indels 155; Gaps 49;  
  
Qy 5 RGLVV---AWALSIMPQFTDTFNMDTKPRVIPSRTAFPGYTVQOCHDISGNKMLVWGP 61  
Db 3 RGVVILLCGWALASCHG---SNLDVEKP-VVFKEADAASFQTVVQF---CGSRLVWGP 54  
Qy 62 LETNGYOKTGDVYKCPVHGNCTKLNAGRTLSNVSRKONMRLGLSLATPKNDSFLAC 121  
Db 55 LEAVAVNQTSQSDCPATGVCQFILL-HIPLEAV-----NMSLGLSLVADTNNSQLAC 108  
Qy 122 SPLMSHBCSGSYTTGMCVRVNSNFRSKTVAPALQRCQ-TYMDIVLVLDGNSI--YPW 178  
Db 109 GPTAQACAKMVAKSGCLLGLSLOQIQAIPATMPCPGQEMDIAFLDGGSGIDQSDP 168  
Qy 179 VEVOHPLINILKXFIYGPQI-----QGVVVOYGEDVHVEHFLNDYRSKDVVZAAASHIE 233  
Db 169 TOMKDFYKALM-----GQLASTSTFSLMQVSNILKTHFTTFEKKSSLSQSLVDAIV 221  
Qy 234 QRGCTETRTAFGIEFARSEAF--QKGRKGAKKVMIVITDCESHDSP-DLEKVTQQSERD 290  
Db 222 QLOGL-TYTASGIQKVVKELFHSKNGARKSAKLLIVITDQKFRDPLEYRHHVPEAKA 280  
Qy 291 NVTRYAVAVLGYNNRRGINPETFLNEIKYTASDDPKHPFNVTDEALKDIDVLDGRIP 350  
Db 281 GIIRYAIQVGDAPRE-----PTALQELNTIGSAPSDHVFQVGNFVALRSIQRIQKIF 335  
Qy 351 SLEGT-NKNETSFGLEMSQTFSSHVVEDGVLLGAVCAVDWNGAVLKETAGKVIPIRES 409  
Db 336 AIEGTERSSSSSQHEMSQEGFSSALSNDGPVLGAVGFGFSWCGAFIYPS-----NMRST 390  
Qy 410 YLKEFPPEELKNHGAYLGTYTTSVVSSRQGRVYVAGAFRNFHTGKVLFTMNNRSLTIHQ 469  
Db 391 FINMSQENEDMRDAYLGYS-TALAFKGVHSLILGAPRHQHTGKVVIPT-QESRHWRFKS 448  
Qy 470 AMRQOQIGSYFGSEITSDIDGGVTDVLLVGAAPMYENEGREGKVVVYELR--ONRFVY 527  
Db 449 EVRGTOIGSYFGASLCSDMDRGGSTDLVLIGPHYEHTR--GGQSVSCMPGVRSRWHC 507  
Qy 528 NGTLKDSHSHYQONARFSSIASVRDLNODSYNDVWVVGAPLENEAGAVIFHG--FRGSILK 586  
Db 508 GTTILGEGQHPGRGAAITVLGDVNGDSLADVAIGAPGEEENRGAIVIFHGASQDIAP 567  
Qy 587 TPQRITASELATGQYFQCSHIGLDLNEGLIDLAVGALGNAMVILMSRPVQINASLH 646

Db 568 SPQRVTGSLFRLQYFGSLGGQDLTQDGLVLAAGQGHVLLRLSLPLKVGISIR 627  
Qy 647 FEPSKINIFHRDCKRGRDATCL---AAFLCFT-----PIFLAPHQTTTIGIRYNATMD 698  
Db 628 FAFSEVAKTYQC--WGRTPVLEAGEATVCLTVKSGSPDLG-----DVQSSVRYDLALD 681  
Qy 699 ERYTTPRAHLDEGDFATRAVLLSSGQELCEINHFVLD--TADYKVPVTFSEVSL--- 754  
Db 682 PGLISRAIPDETNCNLTTRKTLGLGDH--CETMKLLLPDCEVADVTPILRLNLSLAD 740  
Qy 755 EDPDHG--PMLDDGMPITLRSVVPFWNGCNEDEHCVPLDVLDAISDLPTAMEYCORVLRK 812  
Db 741 SAPSRNLRPVLVAGSDHVTASTPPEKNCKQELLCGNL----- 779  
Qy 813 PAQDCSAYTLSTFTTIFIESTRQVAVEMTLENGENAYSTVLNLSQANLOFASLIQK 872  
Db 780 -----GVSNFSGLOVLEVGSSPELTVTVMNDEGDSYGLIKFYPAELSYRVRTA 833  
Qy 873 EDSBG---STECVNE---ERRLOKQVNVSPFPFRAKAVAPRLDSEPSKSIPL-HHLEI 925  
Db 834 QORHPYPLRLACEAETGQESLRSSCSINHPIFREGAKATMITFDVSKAFGLDRLL 893  
Qy 926 ELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFTRSS-----SLSHYEVKLNSSL 977  
Db 894 RASASSENKPEISK--TAPQLELPVKYTYTVVISRQEDSTKHFNFSSHGE--RQKEARH 950  
Qy 978 RYDGIQPPSPICPRIONGLFPIHGMWMTITPIATFSGNRL--LKLRLDFTDEANTSCN 1035  
Db 951 RY-----RVNLSPLTL-AISVNFVPIILL-NGVAVDVTILR---SPAQGVSC- 993  
Qy 1036 IKGNSYETPTPVEEDLRAPQ---LHNSNDVWSINCNI-RLVNPQINPHLLGNL-- 1088  
Db 994 -----VSQREPPQHSLLTQIQRSVLDCAIADCLHRCIDPSLGLTDLDELFIKGNLSF 1048  
Qy 1089 -MLRSL--KALKYKMKVMDALQRFHSPFIFEREEDPSQIEFISQEDWQVPIWI 1144  
Db 1049 GMSQTLQKVKLLSBAETFTNTSVYSQLPQGEAFURAQVSTMEBYVVE-----PVFL 1103  
Qy 1145 IVGSTLGGLLALLVALRKLGFFRSARRRE---PGLDPTP 1184  
Db 1104 MFPSSVGGLLALLIATVLYLGLGFK--ROYKEMLDLPDADPOP 1145

RESULT 6

US-08-482-293A-53  
; Sequence 53, Application US/08482293A  
; Patent No. 5831029  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit  
; CORRESPONDENCE ADDRESS:  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; Zip: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,293A  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 53:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1161 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-482-293A-53  
  
Query Match 19.2%; Score 1194.5; DB 2; Length 1161;  
Best Local Similarity 30.1%; Pred. No. 2.3e-98;  
Matches 374; Conservative 198; Mismatches 507; Indels 165; Gaps 49;  
  
Qy 5 RGLVW--AWALSLMPGFTDTFNDTRKPRVIGSRTAFFGYTVQQHDSIGNKMLVVGAP 61  
Db 3 RGVVILCGWALASCHG---SNLDVEKP-VVPKEDAAAFQGVTVQP---GGSRLVVGAP 54  
Qy 62 LETNGYQXTGVDYKCPVHGNCTKLNLRVTLSVNSERKDMRLGLSLATPKDMSFLAC 121  
Db 55 LEAVAVQQTGSSDCPPATGVCQPIIL-HIPLAV-----NMSLGLSVAGTNNSQLAC 108  
Qy 122 SPLWSHEGSSYTTTGMCSRNVSNFRFSKTVAPALQRCQ--TYMDIVIVLDGSNSI--YPM 178  
Db 109 GPTAQRACAKWYAKGSCLLLGSSLIQIAIPATMPCFCQGEMLAFILDGSGSIDQSDP 168  
Qy 179 VEVOHFLINILKXYIGPGQI-----QGVVVQYGVEDVHFLMDIARSVKDVVGAASHIE 233  
Db 169 TQMKDFVKALM-----GQLASTSTSFSLMOYSNILKTHFTTFEFSKSLSPQSLVDAIV 221  
Qy 234 QRGGTETRTAFGIPIPARSEAF--OKGGRKGAKKVMIVITDGHSDSP--DLEKVIQOOSRD 290  
Db 222 QLOGL--TYTAGIQKVKVKELFHSGKARKSAKILLITDQKXPRDPLEYHVIPEAKA 280  
Qy 291 NVTRYAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIDALGDRIF 350  
Db 281 GIIRYAIGVGDAPRE-----PTALQELNLTIGSAPSQDRVFKVGNFVALRSIQRTQEKIF 335  
Qy 351 SLEGT--KNKETSFGLEMSOTGFSFHHVVDGVLGAVGAYDMGAVLKETSAGKVIPLRES 409  
Db 336 AIEGTESRSSSFQHEMSQBGFSALSMDGVPVLAGVGFSWGGAFLYPS-----NMKST 390  
Qy 410 YLKEFPPEELKNHGAYLGYTVTSVVSSRQGRVYVAGAPRPNHTKGVILFTMNNRSLTIHQ 469  
Db 391 FINMSQENEDMRDAYLGV--TALAFAWKGVHSLILGAPRHQHTGKVWIFT--QESRHWPKS 448  
Qy 470 AMRGOQISYFGSSEITSDIGDGVTVLLVAGPMYNEGERKGVVVELR--QNRVY 527  
Db 449 EVRGTQISYFGSGLCSVDMDRDSGLTLVIGVPHYEHTF--GGQVSCVCPMGVRSRWHC 507  
Qy 528 NGTLKDSHSYQNAFSGSIASVRDLNQPSYNDVVVGVGAPLEDNHAGAIYIFHG--FRGSILK 586  
Db 508 GTTLHGEGQHPWGRFGAALTGLVDVNGDSLADVAIGAPEGSEENRGAVYIFHGASRQDIAP 567  
Qy 587 TPQRITASELATGLQYFGCSIHQQLDINEDGLIDLAVGALGNVILNSRVPVAINSLH 646  
Db 568 SPQRVTGSLFRLQYFGSLGGQDLTQDGLVLAAGQGHVLLRLSLPLKVGISIR 627  
Qy 647 FEPSKINIFHRDCKRGRDATCL---AAFLCFT-----PIFLAPHQTTTIGIRYNATMD 698  
Db 628 FAFSEVAKTYQC--WGRTPVLEAGEATVCLTVKSGSPDLG-----DVQSSVRYDLALD 681

QY 699 ERYTPRAHLDGSGDRFTNRAVLSSQELCERINPHVLD-TADYVVPVTSVEVSL--- 754  
 Db 682 PGLLSRAIFDFTKNTLRTKTLGLGDH-CETMKLLPDCVEDAVTPIILNLNLSLGD 740  
 QY 755 EDPDHG--PMLDDGWPMTLRVSPFWAGCNEDEHCVPLVLDARSDDLPTAMEYQORVLRK 812  
 Db 741 SAPSRNLAPVLAVGSDHVTASFPPEKCKQELLCEGNI----- 779  
 QY 813 PAQDCSAYLSPDTTFFIESTRQRAVEATLENGENAYSTVLNISANLQFASLIQK 872  
 Db 780 -----GVSNFSGLOVLEVGSSPELTVTVTVMNEDGSYGLTKFYFPAELSVRVTRA 833  
 QY 873 EDSDC---SIBCVNE---ERRKQKQCVNVSVPFFRAKAKAFRLDSEPSKSIPL-HHLEI 925  
 Db 834 QOPHPYPLRLACEAPTGOESLRSSCSINHPIFREGAKATFMTDVSYKAFGLDRLIL 893  
 QY 926 ELAAGSDNERSTEDNVAIRFLKYEADVLFTRSS-----SLSHYEVKLNSSLE 977  
 Db 894 RASASSENKPEITSK--TAFQLELPVKYTVTVTVISROBDSKHFNFSSHGE-RQKEAEH 950  
 QY 978 RYDGIQPPFSCIFRIONLGLPFIHGMKMTIPIATRSNRL--LKLKDFLTDEANTSCN 1035  
 Db 951 RY-----RVNLSPLTL-AISNFWPILL-NGVAVMDVTLR---SPAQGVSC- 993  
 QY 1036 IWNSTVPTVPEBDELRAPQ-----LNHSNDVVSINCN-RELVPNOEINFHLLGNL-- 1088  
 Db 994 -----VSQREPPQSHDLTLQIGRSVLDCAIDCLHLRCDIPSLGTLDELDFILKGNLSP 1048  
 QY 1089 -WLRLS---KALKYKSMIMVNAALORQHPFPFREDPSROIEFELISKQEDQVWIWI 1144  
 Db 1049 GWISQTLQKVLLESEBITNTNISVYSQLPQEAFLRAQVSTMLEEYVYV- -PVFL 1103  
 QY 1145 IVGSTGLGLLALLVLAIRKLGFRRSARRRE-----PGLDPTP 1184  
 Db 1104 MVFSSVGLLLALITVALYKLGFFK--ROYKEMLDLPSADPDP 1145

## RESULT 7

US-08-943-363-53  
 ; Sequence 53, Application US/08943363  
 ; Patent No. 5837478  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gallatin, W. Michael  
 ; APPLICANT: Van der Vieren, Monica  
 ; TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit  
 ; NUMBER OF SEQUENCES: 114  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 ; STREET: 233 South Wacker Drive, 6300 Sear Tower  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: United States  
 ; ZIP: 60606-6402  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/943,363  
 ; FILING DATE:  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/173,497  
 ; FILING DATE: 23-DEC-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/286,889  
 ; FILING DATE: 5-AUG-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/362,652  
 ; FILING DATE: 21-DEC-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Williams Jr., Joseph A.

; REGISTRATION NUMBER: 38,659  
 ; REFERENCE/DOCKET NUMBER: 27866/32684  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312-474-6300  
 ; TELEFAX: 312-474-0448  
 ; TELEX: 25-3856  
 ; INFORMATION FOR SEQ ID NO: 53:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1161 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-943-363-53

Query Match 19.2%; Score 1194.5; DB 2; Length 1161;  
 Best Local Similarity 30.1%; Pred. No. 2.3e-98;  
 Matches 374; Conservative 198; Mismatches 507; Indels 165; Gaps 49;

QY 5 RGLVV---ANALSLWPGFTDTENMTKPRVIFGSRTPAGYTVQOHDISGNKWLWVGAP 61  
 Db 3 RGVVILLCGHALASCHG-----SNLDVEKP-VFKEDAAAPGQTVVQF---GGSLVVGAP 54  
 QY 62 LETNGYQKTDGVYKCPVIHGNCTKMLNLRVTLNVSERKDNMRIGLSLATNPKNDSFLAC 121  
 Db 55 LEAVAVNQTCSSDCPEATGVCQPILL-HIPLEAV-----NMSLGLSLVADTNNSQLAC 108  
 QY 122 SPLWSHECGSSYTTGMCVRVNSNFRPSKTVAPALQRCQ-TYNDIVIVLGGNSI--YFW 178  
 Db 109 GPTAQARACAKMYKAGSCLLGLSSLOFIQAIATMPSPCQOEMDIAFLIDGSGSIDQSD 168  
 QY 179 VEYQHFLINILKFFYICPGQI-----QVGVOVQGEDVHVEFHLNDVRSVXDVVEAASHIE 233  
 Db 169 TOMKDFVKALM-----CQLASTSISFSLMQSYNLKTHFTTEPKSSLSQSLVDIV 221  
 QY 234 QRGGTETRTAFGIEPARSEAP--OKGRGAKKMTIVITDGEHSDSP-DLSKVIQOQSERD 290  
 Db 222 QLOGL-TYTAGSIQKVVKELFHSKNGARKSAKILIVITDQKFRDPLEYHRVPIPEAKA 280  
 QY 291 NVTRYAVAVLGYNNRRGINPETFLNRIKYIASDPPDKHPFNVTDEAALKDIDVALGDRIF 350  
 Db 281 GIIRYALGVGDAFRE-----PTALQELNLTIGSAPSQDHVKFKNFVALRSIQRIQEKIP 335  
 QY 351 SLEGT-NKNETSPGLEMSQTGFSSHVVBDGLVGAVGVNDWNGAVLKETSAGKVIPIRES 409  
 Db 336 AIBGTSSRSSSSFQHEMSQEGFSSALSMDBGPVLGAVGVGSWGGATLPS-----NMRST 390  
 QY 410 YLKEFPEELKNHGAVLYGVTVTSVSSRQGRVYVAGAPRNHTGKVLILFTMHNNRSLTIHQ 469  
 Db 391 FIMNSQENEDMDRAYLGIS-TALAFWKGVHSLILGAPRHQHTKVVIFT-QESRHRPKS 448  
 QY 470 AMRGOQIGSYFGSEITSVBIDGQVTDVLLVGAPMYFNEGRERGRKYVYVELR--QNRVY 527  
 Db 449 EVRGTQIGSYFGASLCSVDMDRDGSTDVLVIGVPHYVEHTR--GGQVSVCPMPGVRSRWHC 507  
 QY 528 NGTLKDSHSYQXARFGSSIASVRDLNODSYNDVWVQAPLEDNHAGAIYIFHG-FRGSILK 586  
 Db 508 GTTLHGEQGHWPGRFGAALTIVLGDVNGDSLDVAIIGAPGEEENRGAVYIFHGASRDIA 567  
 QY 587 TPKQRITASELATGLQYFGCSIHGOLDNEDGLIDLAVGALGNVILWSPVVOINASLIH 646  
 Db 568 SPQSRVTSQQLFLRLQYFGCSLGGQDLTQDGLVDLAVGAQGHVLLRLSLPLKLVGISIR 627  
 QY 647 FEPKINIHRDCKKSGRDATCL---AAPLCPT-----PIFLAPHEQTTTGVIRYNATMD 698  
 Db 628 FAPSEVAKTVYQC--WGRTPTVLEAGATVCLTVARKGSPDLLG-----DVQSSVRYDLALD 681  
 QY 699 ERYTPRAHLDGSGDRFTNRAVLSSQELCERINPHVLD-TADYVVPVTSVEVSL--- 754  
 Db 682 PGLLSRAIFDFTKNTLRTKTLGLGDH-CETMKLLPDCVEDAVTPIILNLNLSLGD 740  
 QY 755 EDPDHG--PMLDDGWPMTLRVSPFWAGCNEDEHCVPLVLDARSDDLPTAMEYQORVLRK 812  
 Db 741 SAPSRNLAPVLAVGSDHVTASFPPEKCKQELLCEGNI----- 779

QY 813 PAQDCSAYTILSPDTTFFIIBSTQRVAEATLENROENAYSTVLNISQANLOFASLIQK 872  
 Db 780 -----GVSNFSGLOVLEVGSSPELTVTVTWNEGEDSYGLIKFYPAELSYRVRTRA 833  
 QY 873 EDSG---STECVNE---ERRLOKQVCNVSYPFRAKAKVAFRLDSEPSKSIPL-HHLEI 925  
 Db 834 QQHPYPLRLACEAETGOESLESSCSINHPIFREGAKATMTITFDVSYKAFGLDRLLL 893  
 QY 926 ELAAGSDNERSSTKEDNVAPLRFHLKYEADVLFTRSS-----SLSHYEVKLNSSLE 977  
 Db 894 RASASENNKPKTSK--TAPQLELPVKYTVTVVISRQEDSTKHFNFSHSGE--RQKEAEH 950  
 QY 978 RYDGIGPPFSCIFRIQNLGLFPIHGMWMTITPIATRSNRL--LKLRFDLTDEANTSCN 1035  
 Db 951 RY-----RVNLSPLTL-AISVNFVWPILL-NGVAVDVTILR-----SPAQGVSC- 993  
 QY 1036 IWGNSTEYRPTVEEDLRAPO-----LNHNSDVSINCNI-RLVNPQEIHFLLGNL-- 1088  
 Db 994 -----VSQREPPQHSLLTQIQGRSVLDCAIDCLHLRCDIPSLGTLDELDFILKGNLSF 1048  
 QY 1089 -WLRLS---KALKYKSMKIMVNAALQRPSPHPIFREDDPSQIPEISKQEDWQVPIWI 1144  
 Db 1049 GWISQTLQKVKVLLLSABEITFTSVYSQLPQGEAFILRAQVSTMLESYVYVE-----PVFL 1103  
 QY 1145 IVGSTLGGILLALLVLALRLKLGPPFRSARRRE-----PGLDPTP 1184  
 Db 1104 MVFSSVGGILLALLITVALYKLGFFK--ROYKEMLDLPSADPDP 1145

RESULT 8  
 US-09-193-043-53  
 ; Sequence 53, Application US/09193043  
 ; Patent No. 6251395  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gallatin, Michael W.  
 ; TITLE OF INVENTION: No. 6251395el Human 2  
 ; FILE REFERENCE: 27866/35004  
 ; CURRENT APPLICATION NUMBER: US/09/193,043  
 ; CURRENT FILING DATE: 1998-11-16  
 ; EARLIER APPLICATION NUMBER: 08/173,497  
 ; EARLIER FILING DATE: 1993-12-23  
 ; EARLIER APPLICATION NUMBER: 08/286,889  
 ; EARLIER FILING DATE: 1994-08-05  
 ; EARLIER APPLICATION NUMBER: 08/362,852  
 ; EARLIER FILING DATE: 1994-12-21  
 ; EARLIER APPLICATION NUMBER: 08/943,363  
 ; EARLIER FILING DATE: 1997-10-03  
 ; NUMBER OF SEQ ID NOS: 114  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 53  
 ; LENGTH: 1161  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-09-193-043-53

Query Match 19.2%; Score 1194.5; DB 3; Length 1161;  
 Best Local Similarity 30.1%; Pred No. 2.3e-98;  
 Matches 374; Conservative 198; Mismatches 507; Indels 165; Gaps 49;  
 QY 5 RGLVV---ANALSWPGFTDFNMDTRKPRVPGSRPTAFGTGYVQOHDISGNKWLWVGAP 61  
 Db 3 RGWVILCGWALASCHG---SNLDVEXP-VVFKEAASFGQTVVQF---GGSRLVVGAP 54  
 QY 62 LETNGYQKTDGVYKCPVTHGNKTNLGRVTLNSVSEKNNRNLGLSLATNPKDNLPLAC 121  
 Db 55 LEAVAVNOTGSSDCPPATGVCQPIILL-HIPEAV-----NMSLGLSLVADTNNSQLLAC 108  
 QY 122 SPLMSHEGSSYYITGMCVRNSNFRFSKTVAPALQRCQ--TYMDIVIVLQGSNSI--YPW 178  
 Db 109 OPTAQRACAKNMYAKSGCLLIGSSLOFTQOIPATWPCPGQEMDIAFLIDGSGSIDQSDF 168

QY 179 VEQHFELINILKXFIYIGPQI-----QGVVVOYGEDVWHEFLNDYRSKVDVWEAASHIE 233  
 Db 169 TOMKDFVKALM-----GQLASTSTGFSLMQYSNILKTHFTTFTPEKSSLSQSLVDIAV 221  
 QY 234 QRGTEHTRTAFGIBPAREAP--QKGGKKGAKKMWITDGHSDSP--OLEKVIIOOSRD 290  
 Db 222 QLOGU-TYTAGGIQVVKELFHSKNGARKSAKKILIITDQCKPRDPLUETHYRHVIPAESA 280  
 QY 291 NVTRYAVAVLGYVNRGINPETFLNEIKYIASDPDDKGFNFVNTDEAALKDIDVALGDRIF 350  
 Db 281 GIIRYALGVGDAPRE-----PTALQELNTIGSAPSQDHVFKVGNFVALRSIQRIQEKIF 335  
 QY 351 SLEGT-NKNETSGLEMSQCFSSHVEDGVILGANGAYDWNGAVILKETSAGKVIPIRES 409  
 Db 336 AIBETESRSSSPHEMSQEGFSALSNDGPVLGAVGFGSSGAGFYVPS-----NMKST 390  
 QY 410 YLKEFPPELKNHGYLGYTIVTVSVSSRQGRVYVAGAPRFNHTGKRVILFTMHNNRSLTIHQ 469  
 Db 391 FINMSQENEDMRDAYLGS-TALAFWKGVHSLILGAPRHQHTGKVIFT--QESSHWPKS 448  
 QY 470 AMRCQOIGSYGSEITSDIDGCVTVLLVAGPMYENEGREKQVTVYELR--QNRPTY 527  
 Db 449 EVRGTQISYFGASLCSDMDRDGSTDVLIGVPHYEYHTR--GGQVSVCEMPGVRSWHC 507  
 QY 528 NGTLKDSHSYQNAFSGSIASVRDLNODSYNDVVVVGAPLEDNHAGAIYIFHG-FRGSILK 586  
 Db 508 GTTLHGEQGHWPGRFGAALTVLGDVNGDSLADVAIGAPGEENREGAVYIFHGASRQDIAP 567  
 QY 587 TPQRITASBLATGLQYFGCSIHGOLDLNEDELIDLAGALGNVILWSRPPVQINASLH 646  
 Db 568 SPSQRTVSGQLFRLQYFGQSLSGGQDLTQDGLVDLAVGAGQHVLLRLSLPLKVGISIR 627  
 QY 647 PEPKINIFHRDCKRGRDATCL---AALFCFT-----PIFLAPHFQTTVTGIRYNATMD 698  
 Db 628 PAFSEVAKTVQC--WGRTPTVLEAGETVCLTVRKGSPDLG---DVQSSVRYDLALD 681  
 QY 699 ERRYTPRAHIDEGDRETNRAVLLSSQOELCERINFHVL-DADYVVPVTFSEYSL---754  
 Db 682 PGLRIGSAIFDETNCNCTLRRKTLGLGDH-CETMKLLLPDCVEDAVTPIILRLSLAGD 740  
 QY 755 EDPDHG--PMLDDGWPTTLRVSVFPWNGCNEDEHCVPLDVLDAESDLPAMEYCORVLR 812  
 Db 741 SPSRNLRLPVLA VGSQDHVTASTPFPEKNCKQELLCEGNL-----779  
 QY 813 PAQDCSAYTILSPDTTFFIIBSTQRVAEATLENROENAYSTVLNISQANLOFASLIQK 872  
 Db 780 -----GVSNFSGLOVLEVGSSPELTVTVTWNEGEDSYGLIKFYPAELSYRVRTRA 833  
 QY 873 EDSG---STECVNE---ERRLOKQVCNVSYPFRAKAKVAFRLDSEPSKSIPL-HHLEI 925  
 Db 834 QQHPYPLRLACEAETGOESLESSCSINHPIFREGAKATMTITFDVSYKAFGLDRLLL 893  
 QY 926 ELAAGSDNERSSTKEDNVAPLRFHLKYEADVLFTRSS-----SLSHYEVKLNSSLE 977  
 Db 894 RASASENNKPKTSK--TAPQLELPVKYTVTVVISRQEDSTKHFNFSHSGE--RQKEAEH 950  
 QY 978 RYDGIGPPFSCIFRIQNLGLFPIHGMWMTITPIATRSNRL--LKLRFDLTDEANTSCN 1035  
 Db 951 RY-----RVNLSPLTL-AISVNFVWPILL-NGVAVDVTILR-----SPAQGVSC- 993  
 QY 1036 IWGNSTEYRPTVEEDLRAPO-----LNHNSDVSINCNI-RLVNPQEIHFLLGNL-- 1088  
 Db 994 -----VSQREPPQHSLLTQIQGRSVLDCAIDCLHLRCDIPSLGTLDELDFILKGNLSF 1048  
 QY 1089 -WLRLS---KALKYKSMKIMVNAALQRPSPHPIFREDDPSQIPEISKQEDWQVPIWI 1144  
 Db 1049 GWISQTLQKVKVLLLSABEITFTSVYSQLPQGEAFILRAQVSTMLESYVYVE-----PVFL 1103  
 QY 1145 IVGSTLGGILLALLVLALRLKLGPPFRSARRRE-----PGLDPTP 1184  
 Db 1104 MVFSSVGGILLALLITVALYKLGFFK--ROYKEMLDLPSADPDP 1145

## RESULT 9

US-09-688-307A-53  
; Sequence 53, Application US/09688307A  
; Patent No. 6432404  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 6432404el Human Beta-2  
; FILE REFERENCE: 27866/36646  
; CURRENT APPLICATION NUMBER: US/09/688,307A  
; CURRENT FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 09/193,043  
; PRIOR FILING DATE: 1998-11-16  
; PRIOR APPLICATION NUMBER: 08/605,672  
; PRIOR FILING DATE: 1996-02-22  
; PRIOR APPLICATION NUMBER: 08/173,497  
; PRIOR FILING DATE: 1993-12-23  
; PRIOR APPLICATION NUMBER: 08/286,889  
; PRIOR FILING DATE: 1994-08-05  
; PRIOR APPLICATION NUMBER: 08/362,652  
; PRIOR FILING DATE: 1994-12-21  
; PRIOR APPLICATION NUMBER: 08/943,363  
; PRIOR FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 53  
; LENGTH: 1161  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-688-307A-53

Query Match 19.2%; Score 1194.5; DB 4; Length 1161;  
Best Local Similarity 30.1%; Pred. No. 2.3e-98;  
Matches 374; Conservative 198; Mismatches 507; Indels 165; Gaps 49;

QY 5 RGLVV---AWALSLWPGTDFNMDTRKPRVPGSRATFGYTVQOHDISNKNLWVGAP 61  
DB 3 RGVILLCCNALACHG-----SNLDVEKP-VVFEDAAAFQGVVQF---CGSLVVGAP 54  
QY 62 LETNGYQKTDVYKCPVHNGCTKLNLRVTLNVSERKNRKLGLSLATNPKDMSFLAC 121  
DB 55 LEAVAVNQTQSSDPPATGVCQPIIL-HIPLEAV-----NNSLGLSLVADPTNNSQLLAC 108  
QY 122 SPLSHCEGSSYYTTCMSRVNSFRSKTVAPLQRCQ-TYMDIVIVLDSNSI--YPW 178  
DB 109 GPTAQRACAKMYAKGSCLLGSSLQFTQAIPTATMPCPGQEMDIAFLIDGSGSIDQSF 168  
QY 175 VEVOHFLINILKFFYIGPGQI-----QGVVQVQGEDVVHFLNDRYSRVDVVEAASHIE 233  
DB 169 TQKDFVKALM-----GQLASTSTFSLMQYSNLIKTHFTTPEKSSLSPQSLVDAIV 221  
QY 234 QRGGTETRTARGIEFAKSEAF--QKGRKGAKKVMIVITGESHDSP-DLEKVIQOGERD 290  
DB 222 QLOGL-TVTASGIQKVKELPHSKNGARKSAKKILIVITDQKFRDPLEYRHVPEAKA 280  
QY 291 NVTYAVAVLGYNNRGINPETFLNEIKYIASDDPKDFNVNVTDEAALKDIDVALGRIF 350  
DB 281 GIIRYAVGVDAFRE-----PTALQELNTIGSAPSDHVFKVGNFVALRSIQRIQKIP 335  
QY 351 SLRET-KNNETSPGLEMSQTGFSHVVEDGVLLGAVGAYDNMGAVLKETSAGKVIPIRES 409  
DB 336 AIEGTESRSSSPCHEMSQEGFSSALMDGCVLGAVGGSFGSAGAFLYPS-----NMRST 390  
QY 410 YLKEFPBELKNHGLVGYTVTVSVVSRQGRVYVAGAPREHVTGKVLFTMHNRSLIHQ 469  
DB 391 FINMSQENEDRDYLGYS-TALAFKVGWLSLILGAPRHQHTGKVIFT-QESGRHWRPKS 448  
QY 470 AMRQQIGSYRGSBITSDIDGDGVTDLVLLGAFMYFNEGRERKGVVYELR--QNRVY 527  
DB 449 EVRGTQIGSYFGALSCLSDVMDRDGSTDVLVIGVPHYVEHTR-GQGVSVCPMPGVRSRWHC 507  
QY 528 NGTLKDSHSYQARFGSISASVRLNODSYNDVWVGAPLEDNHAGAIYIIFHG-PRGSILK 586

508 GTTLHCEGHPWGRFGAALTIVLGVNDSLADVAIGAAGBEENGAVYIFHCASRQDIAP 567  
587 TPQORITASELATGLQYFGCSIHQDLNEDGLIDLAVGALGNVILWSRPPVQINASLH 646  
568 SPSQRTGSQLFLRLQYFGQSLGGQDLTQDGLVDLAVGAGCHVLLRLSLPLLVGISIR 627  
647 FEPSKINIPHRDCKRGRDATCL--AAFLCFT-----PIFLAPHFQTTTGVIRYNATMD 698  
628 FAFSEVAKTVQC--WGRTPVLEAGEATVCLTVRKGSPLL--DVGSSVRYDIALD 681  
699 ERRYTPRAHLDEGGDFRTRNRAVLLSSQOELCERINFHVL-DADYVVPVTFVSYSYL--- 754  
682 PGRLLISRAIFDETKNCTLTRKTLGLGDH-CETWKLILLPCVEDAVTPIILRLNLSLAD 740  
755 EDPDHG--PMLDDGMPPTLVSVFPGWNGNEDEHCVPLDVLDAARSOLDPTAMEYCORVLR 812  
741 SPSRNLRPVLVGSQDHVTASPPFEKNCQOELLCEGNL----- 779  
813 PAQDCSAYTLTSDFTTPIIESTQRVAVEATLENRGENAYSTVINISQSANLOFASLIQK 872  
780 -----GVSNFSLQVLEVGSSPELTVTVTVANEGEDSYGLTIKEYYPALSVYRVTRA 833  
873 EDSDG---SIECVNE---ERRLOQVCNVSYPFRAKAKVAFRLDSEFSKIFL-HHLEI 925  
834 QQPEPYPLRLACEAEPTEQESLSSSCSINHPFREGAKATFMTFDVSVYKAFGLGDRLL 893  
926 ELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFTRSS-----SLSHVEVKLNSLE 977  
894 RASASSENKPETSK--TAPQLELPVKYTVTVVISRQEDSTKHFNFSSSHCB-RQKEAEH 950  
978 RYDGIOPPPFCIFRIQNLGLFPIHGMKMTIPIATRSNRL--LKURDLTDEANTSCN 1035  
951 RY-----RVNLSPLTL-AISVNFVWPILL-NGVAVMDVTIL--SPAQGVSC- 993  
1036 IWGNSTEYRTPVEEDLRRAPQ-----LNHSNSDVWSINCNI-RLVPMQENFHLGCL-- 1088  
994 -----VSQREPPHSDLLTQIQRSVLDCAICHLRCDIPSLGTLDELDFILKGNLSP 1048  
1089 -WLRSL---XALKYKSKIMVNAALQRFHSPIFREEDPSRQIEFEISKQEDQVPIWI 1144  
1049 GWISQTLQKXVLLSEAITFNTSVSOLPGQAFRLAQVSTMLEBYVVE-----PVFL 1103  
1145 IVCSTIGGILLALLVLRKLGKGFPSARRE-----PGLDPTP 1184  
1104 MFSVSGVGLLLALITVALYKLGFFK--RQYKEMLDLPSADPP 1145

RESULT 10  
US-09-350-259-53  
; Sequence 53, Application US/09350259  
; Patent No. 6620915  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 6620915el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350,259  
; CURRENT FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: 09/193,043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 53  
; LENGTH: 1161  
; TYPE: PRT

Db 894 RASASENNKPKTSK--TAFQLELPVKYTVTVTVISROBDSKGFHFNSSHGE-RQKAEH 950  
 QY 978 RYDGICPPFCIFRIQNLGLPFIHGMMKTIPIATRSNRL--LKLRLDFTLDEANTSCN 1035  
 Db 951 RY-----RNNLSPLTL-AISVNFVWPILL-NGVAVWDVTLR---SPAQGVSC- 993  
 QY 1036 IWGNTSTVRPTPVEEDLBRAPQ----LNHSNSDVVSNCNI-RLVPNQEIHFLLGL-- 1088  
 Db 994 -----VSOREPPQHSDDLTOQGRSVLDCAIADCHLRCDIPSLGTLDLDFILKXNLSF 1048  
 QY 1089 -WLRLS--KALKYKSMKIMVNAALQROFHSPPFIREDPSRQLEFEISKQEDMQVPIWI 1144  
 Db 1049 GWSOTLQKVKLLSEAEITNTSVYSQLPQCEAFLAQAQVSTMLEEYVYVY- - - - -PVFL 1103  
 QY 1145 IVGSLTGLGLLIALILVLAIRKLGFPFSARRRR- - - - -PGLDPTP 1184  
 Db 1104 MYPSSVGGELLIALITVALYKLGFFK--ROYKEMLDLPSADPDP 1145  
 RESULT 11  
 US-09-193-043-55  
 ; Sequence 55, Application US/09193043  
 ; Patent No. 6251395  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gallatin, Michael W.  
 ; APPLICANT: Van der Vieren, Monica  
 ; TITLE OF INVENTION: No. 6251395el Human 2  
 ; FILE REFERENCE: 27866/35004  
 ; CURRENT APPLICATION NUMBER: US/09/193,043  
 ; EARLIER FILING DATE: 1998-11-16  
 ; EARLIER APPLICATION NUMBER: 08/173,497  
 ; EARLIER FILING DATE: 1993-12-23  
 ; EARLIER APPLICATION NUMBER: 08/286,889  
 ; EARLIER FILING DATE: 1994-08-05  
 ; EARLIER APPLICATION NUMBER: 08/362,652  
 ; EARLIER FILING DATE: 1994-12-21  
 ; EARLIER APPLICATION NUMBER: 08/943,363  
 ; EARLIER FILING DATE: 1997-10-03  
 ; NUMBER OF SEQ ID NOS: 114  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 55  
 ; TYPE: PRT  
 ; LENGTH: 1161  
 ; ORGANISM: Rattus rattus  
 ; ORGANISM: Rattus rattus  
 US-09-193-043-55  
 Query Match 19.0%; Score 1184.5; DB 3; Length 1161;  
 Best Local Similarity 29.6%; Pred. No. 1.8e-97;  
 Matches 361; Conservative 207; Mismatches 519; Indels 131; Gaps 43;  
 QY 7 LVVAWALSFWPGFTDTFMMDTRKPRVPGSRTAFPGYTVQOHDISGNKLVVGAFLPTNG 66  
 Db 8 LLCGWAVLASCHG---SNLDVEEP-IVFREDAASFGQTVQF---GSRLLVVGAPLEAVA 59  
 QY 67 YQGTGVYKCPVIEHCNCTKLNGLVTLNSYSEKDNMRGLSLATNPKNSFLACSPWS 126  
 Db 60 VNQGRYLDCAPATGMCQPIVL-RSPLEAV-----NMSLGLSLVATNNAQLLACGPTAQ 113  
 QY 127 HECGSSYTTGACSRVNSNFRFSKTVAPALQRC-QTYMDIVIVLDGNSNI--YPMVSVQH 183  
 Db 114 RACVKNMAYKSGCLLLGSSLOFIAQVAPASMPCECPQEMDIAFLIDGSGSINORDFAQMKD 173  
 QY 184 FLINILKPYIGPGQIQVGVVQYGEDVVHFEHLMDYRSKDVVEAASHIBORGCTERTTA 243  
 Db 174 FYKALMGFF--ASTSLFSLMCMYSNILKTHFTTFEFKNILDPQSLVDPIVOLQSL-TYTA 230  
 QY 244 FGIEEFARSEAF--QKGRGKAGKVMIVITDGESHDP-DLEKVTQOSRONVTRYAVAVL 300  
 Db 231 TGIRTVMEELFHSKNGSKSAKILLVITDQKYRDPDLEYSDVTPADKAGIIRYALGVG 290  
 QY 301 GYNRRGINPETFLEINIKYIASDDDDKHFFNVTDAAALKOIVDALGRIFSELEGT--NNNE 359  
 Db 291 DAFQE-----PTALKELNTIGSAPPQDHVFKVGNFAALRSIQROLOQKIFAIEGTQSRSS 345  
 ; ORGANISM: Mus musculus  
 US-09-350-259-53  
 Query Match 19.2%; Score 1194.5; DB 4; Length 1161;  
 Best Local Similarity 30.1%; Pred. No. 2.3e-98;  
 Matches 374; Conservative 198; Mismatches 507; Indels 165; Gaps 49;  
 QY 5 RGLVV--AAWALSFWPGFTDTFMMDTRKPRVPGSRTAFPGYTVQOHDISGNKLVVGAFL 61  
 Db 3 RGVVILCCGVALASCHG-----SNLDVEEP-VFKEDAAASFGQTVQF---GGSRLVVGAP 54  
 QY 62 LETNGYQKTDGVYKCPVIEHCNCTKLNGLVTLNSYSEKDNMRGLSLATNPKNSFLAC 121  
 Db 55 LEAVAVNQTGSSDCEPPATGVCPILL-HIPLAV-----NMSLGLSLVADTNNSQLLAC 108  
 QY 122 SPLMSHECGSSYTTGACSRVNSNFRFSKTVAPALQRC-QTYMDIVIVLDGNSNI--YPM 178  
 Db 109 GPTAGACAKMAYKSGCLLLGSSLOFIAQVAPASMPCECPQEMDIAFLIDGSGSIDQSD 168  
 QY 179 VEVQFLLINILKPYIGPGQI-----QVGVQYGEDVVHFEHLMDYRSKDVVEAASHIE 233  
 Db 169 TQMKDFVKALM-----COLASTSTSLMCMYSNILKTHFTTFEFKSSLSPLQSLVDAIV 221  
 QY 234 QRGGTETRTAFGIEFARSEAF--QKGRGKAGKVMIVITDGESHDP-DLEKVTQOSERD 290  
 Db 222 QLQGL-TYTAGIQKVKELFHSKNGSKSAKILLVITDQKYRDPDLEYSRVHVIPEAKA 280  
 QY 291 NVTREAVAVLGYNRRGINPETFLEINIKYIASDDDDKHFFNVTDAAALKOIVDALGRIF 350  
 Db 281 GIIRTAIVGDAFRE-----PTALQELNTIGSAPPQDHVFKVGNFAALRSIQROIQBKIF 335  
 QY 351 SLEGT-NKNETFGLMSQTFSSHHVDEGVLGAVGAYDMNGAVLKETSACKVIPERES 409  
 Db 336 AIEGTESRSSSFQHEMSQEGFSSALSMDGPIVLAGVGGFSSGGAFLYPS-----NKST 390  
 QY 410 YLKEPPEELKNHGAIVGTVTVSVSRQGRVTVAGAPRNTGKVLFTMNNRSLTIHQ 469  
 Db 391 FIMSQENEDMDAYLGVS-TALAPFKGWHSILGAPRHOHTGKVIPT-QESRHWPKS 448  
 QY 470 AMRGOQIGSYFGSEITSVDIDGQVTDVLLVGAQPMYFNEGRERGKVVYVELR--ONRFVY 527  
 Db 449 EYRGTOIGSYFGASLCSVMDRDGSTDVLIGVPHYEYTR-GQVSVCPMPGVRSRWHC 507  
 QY 528 NGTLKDSHGYQNAFGSSIASVRDLNQSYNDVVGAPLEDNHAGAIYIFHG-FRGSILK 586  
 Db 508 GTTLHGEOGHPWGRFGAALTVLGDNVGSADVAIGAPGEBENRGAVYIFHGASRQDIAP 567  
 QY 587 TPQRITASELATGLOYFCSSHGGOLDLNEOGLDLAVGALGNVILKRSVPVVOINASLH 646  
 Db 568 SPQSRVTGSLFLRLOYPQGLSGQODLTQGLVDLAVGAQHLLRSLELLKVGISIR 627  
 QY 647 FEPSKINIHRDKRSGRDATCL--AAFLCFT-----PIFLAPHFQTTTVGIRYNATMD 698  
 Db 628 FAPSEVAKTVYQC--WGRTPTVLEAGATVCLTVRKGPSDLLG---DVQSSVRYDIALD 681  
 QY 699 ERYTPRAHLDEGDEFTNRAVLSSGQELCERINPHVLD-TADYVKPVPFSVEYSL--- 754  
 Db 682 PGLISRAIFDTKNTCLTRKTLGLGDH-CETMKLLPDCVEDAVTPPIHLRLMLSLAGD 740  
 QY 755 EDPDHC--PMLDDGWPTTLRVSPVPFWNGCNEHCVPLDLVLDARSDLPTAMEYQORVLK 812  
 Db 741 SAPSRNLRPVLAGSQDHVTAFFPEKKNCKQELLCEGNL----- 779  
 QY 813 PAQDCSAYTSLSDTTFVIESTRQRAVEATLENGENAYSTVLNISQANLQFASLIQK 872  
 Db 780 -----GVSNFNSGLQVLEVGSSPLBTVTVTWNNEGEDSYGLTKFYYPBELSVRRVTRA 833  
 QY 873 ESDSG--SIECVNE--ERRLOKQVCNVSVPFFPRAKAKVAFRLDSFESKSIPL-HLEI 925  
 Db 834 QQHPVPYPLACEAPTQGESLRSSCSINHPIFREGAKATFMTTFDVSVYKAPLGRDLL 893  
 QY 926 ELAAGSDNERSDXTEDNVAPLRFHLKYEADVLFTRSS-----SLSHVEYKVLNSSL 977

```

QY 360 TSPGLEMSQTGSSSHVVDGVLGAVGAYDNGAVLKETSAGKVP--LRSEYLFKEPPE 417
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
346 SSFQHEMSQEGFSSALTSQGLVAVGVSFGSGGAP-----LYPNTTRPTFFINMSQEN 398
QY 418 LKHGAYLGYTTSVVSQRQVYVAGAPRNFHTCKVILFTWNNRSLTIHQARGGQIG 477
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
399 VDMRDSYLGYS-TAVAFWKGVHSLILGAPRHCHTKVIFT-QEAWHRPKSEVRGTQIG 456
QY 478 SYFGSEIISVDIDGVDGTVLLVGAAPMVFNEGRERKGVYVELR--QNRVYNGTLKDSH 535
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
457 SYFGASLCSVDVDRDGSXDLVLIGAPHYYEQTR--GGQVSVPFVGVGRWQCEATLHGQ 515
QY 536 SYQNARFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHG-FRGSILKTPKQIRTA 594
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
516 GHPWGRFVALTVLGDVNGDNLADVAIGARBEESRGAVIFHGASRLIEIMPSPSRQV 575
QY 595 SELATGLQYFGCSIHGQJDLNEDGLIDLAVAGLGNVILMRPVPVQINASLHFPSPKINI 654
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
576 SGLSLRLQYFGQSLSGGDLTQDGLVLDVAVGAQGHVLLSLPLKLVLSIRFAPMEYAK 635
QY 655 PHRDCKRSGRDATCL--AAPLCFTPIFLAPHPTTTVG-IRYNATMBERYTTPRAHLD 710
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
636 AVYQWE--RPTTVLEAGEATVCLTVHKGSDPLLGNVQGSVRYDLALDPGLISRAIFDE 693
QY 711 GGDRETNRAVLSSQCELCERINFHVLDTADYVKPVTFVSVEYSI-----BDPDHGMKLD 764
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
694 TANKTLTGKTKLGLGDH-CETVKLLPDCVEDAVSPIILRLNLFSLVRDSASPRNLHPVLA 752
QY 765 DGPWTLVSVVPFNGKNEDEHCVPLVDLARSDLPTAMEYQVRVLRKPAQCCSAYTLSP 824
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
753 VGSQDHITASLGFPEKCKQELLCEGDL-----GISENF 785
QY 825 DTTVFIIESTORVAEATLENRGENAYSTVLNISQSANLQFASLI-----QKEDSDGSEIC 881
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
786 SGLQVLVVGSGSPELTVTVTWNEGDSYGLVKFVYPAGLSVRRVTGTQPHQVPLRLAC 845
QY 882 VNE-----ERRLQKQCNVSPFPRAKAKVAFRLDSEFSKIFL-HHELEIELAAGSDSNERD 937
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
846 EAPPAQEDLSSSCSINHIFREGAKTFTMTFVSYKAFGLDRLLRAXASBENKPD 905
QY 938 STKEDNAPLRFHLYEADVLFTRSSSLSHVEKLNLSLERYDVGICPPFCIFRQNLGL 997
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
906 TNK--TAQLELPVKYTVYLLISQEDSTNH-VNFSSS--HGGRQEAARHYRVNLSLP 959
QY 998 PIHGMKKITIPATRSNRLKURDLTDBANTSCNMGNSYETPTVEE-----DL 1052
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
960 LKL-AVRVFWFVPL-----LNGVAVMDVTLSSPAQGVSCVSKPPQNPEDFLTQI 1009
QY 1053 REAPOLNHSNDVVSINCNIRLVNQ-EINPHELGNL---WLSLI---KALKYKSMKTMV 1105
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1010 QRSVLDCSIADCLHFRCDIPSLDIQDELDFILRGNLSFGWVSQTLQEKVLLVSEAITF 1069
QY 1106 NAALQRFHSPPIFREDDPSRQIFESKQEDWQV--PIWTVSTGLGLLALLVALL 1163
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1070 DTSVYSQLPQGAFLR---AQVETL---BEYVYVEDIFLVAGSVVGLLALLITVVL 1122
QY 1164 RKLGFRRARRRREPGLD 1181
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1123 YKLGFP---XKROYKEMLD 1137

```

RESULT 12  
 US-09-688-307A-55  
 ; Sequence 55, Application US/09688307A  
 ; Patent No. 6432404  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gallatin, Michael W.  
 ; APPLICANT: Van der Vieren, Monica  
 ; TITLE OF INVENTION: No. 6432404el Human Beta-2  
 ; FILE REFERENCE: 27866/36646  
 ; CURRENT APPLICATION NUMBER: US/09/688,307A  
 ; CURRENT FILING DATE: 2000-10-13

```

; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605,672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173,497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Rattus rattus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 474
; OTHER INFORMATION: Xaa = any or unknown amino acid
; NAME/KEY: misc_feature
; LOCATION: 1128
; OTHER INFORMATION: Xaa = any or unknown amino acid
; US-09-688-307A-55

Query Match          19.0%; Score 1184.5; DB 4; Length 1161;
Best Local Similarity 29.6%; Pred. No. 1.8e-97;
Matches 361; Conservative 207; Mismatches 519; Indels 131; Gaps 43;

QY 7 LVAVALLSLMPGTTDTFMDTRKPRVIFGSRTPAPGVTVQCHDISGNKWLWVGAPLEFNG 66
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 LLCGWVLASCHG---SNLDVEEP-IVFREDAASFGQTVVQF---GGSRLLVVGAPLEAVA 59
QY 67 YQKGVGVKCPVTHGCTKLNGLRVTLNVSRKDNRLGLSLATNPKNSPLACSLPLWS 126
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
60 VNQGRLYDCAPATGMCQPIVL-RSPLEAV-----NMSLGLSLVATNNAQLLACGPTAQ 113
QY 127 HECSSYTTTGMCSRVNSNFRPSKTVAPALQRC-QTYMDIVLVDGNSI--YPMVEVOH 183
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
114 RACVKNMYAKGSCLLLGSSLQFIQAVPASPMPCEPQROEMIAFLTDGSGSINQORDFAQMD 173
QY 184 FLNLLKKEVIFGCIQGVGVQYGEDVVFHFLNDVSVKDVVEAAASHIQRGGTETRTA 243
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
174 FVKALMGEP--ASTSLFSLMQYSNLLKTHFTTEFKNILDPSQSLVDPIVQLQGL-TYTA 230
QY 244 FGIEFARSEAF--QKGRKGAKKVMYITDGEHSDSP-DLEKVIQOESERNVTRYAVAVL 300
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
231 TGIRTVMEEELFHSGKSGSKSAKILLVITDQKYRDEPLEYSDVIPAADKAGILRYAIGVG 290
QY 301 GYNRRGINPTFLNEIKYIASDDDKHPFNVTDEAKLDIVDALGDRIFSLGCT-NKQE 359
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
291 DAFQSE-----PTALKELNTTIGSAPPQDHVFKVGNFAALRSIQRLQSEKIFAIBGTQSRSS 345
QY 360 TSPGLEMSQTGSSSHVVDGVLGAVGAYDNGAVLKETSAGKVP--LRSEYLFKEPPE 417
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
346 SSFQHEMSQEGFSSALTSQGLVAVGVSFGSGGAP-----LYPNTTRPTFFINMSQEN 398
QY 418 LKHGAYLGYTTSVVSQRQVYVAGAPRNFHTCKVILFTWNNRSLTIHQARGGQIG 477
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
399 VDMRDSYLGYS-TAVAFWKGVHSLILGAPRHCHTKVIFT-QEAWHRPKSEVRGTQIG 456
QY 478 SYFGSEIISVDIDGVDGTVLLVGAAPMVFNEGRERKGVYVELR--QNRVYNGTLKDSH 535
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
457 SYFGASLCSVDVDRDGSXDLVLIGAPHYYEQTR--GGQVSVPFVGVGRWQCEATLHGQ 515
QY 536 SYQNARFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHG-FRGSILKTPKQIRTA 594
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
516 GHPWGRFVALTVLGDVNGDNLADVAIGARBEESRGAVIFHGASRLIEIMPSPSRQV 575
QY 595 SELATGLQYFGCSIHGQJDLNEDGLIDLAVAGLGNVILMRPVPVQINASLHFPSPKINI 654

```



Db 576 SLSLRLQYFGQSLGGQDITQDGLVLAQAGHVLRLSLPLKVKELSIKPAPEVAK 635  
 Qy 655 FHRDCRSGRDATCL---AAFLCFTPIFLAPHFQTTTGV-IRYNATMDERRYPRAHLDE 710  
 Db 636 AVQCWE--RTPTVLEAGEATVCLTVHKSGPDLGNGVQSVRYDLALDPCRLISRAIFDE 693  
 Qy 711 GDRFTNRAVLSSGOELCERINPHVLD-TADYVKPVTFSVEYSL-----EDPDHGMPLD 764  
 Db 694 TKNCTLTGRKTLGLGDH-CETVKLLLPDCEVEDAVSPIILRLNLSLRVDSASPNLHPVLA 752  
 Qy 765 DGNPTTLRVSPVFWNGCNDEHCVDPDLVDARSDELPTAMEYCORVLRKPAQDCSAYTLSP 824  
 Db 753 VGSQDHIATSLPFKNCKQELCEGDL-----GISFNF 785  
 Qy 825 DTTVTIESTRQVAVATLENGENAYSTVLNISQSANLQFASLI---QKSDSGSIEC 881  
 Db 786 SGLQVLVVGSPSLTIVTVWNEGEDSYGLTVKFPYFAGLSYRVGTQOHPYPLRLAC 845  
 Qy 882 VNE---ERRLQKQVNCVSYFFRAKAKAVAFRLDSEFSKIFL-HHLEIELAAGSDNERD 937  
 Db 846 EAEPAQEDLRSSSCSINHPIFREGAKTTFWITFDVSKAFGLDRLLRKAKASSENKPD 905  
 Qy 938 STKEDNVAPLRFLKYRADVLFTRSSLSHYEVLKNSLERYDGIOPPFSCIFRIONLGL 997  
 Db 906 TNK--TAFQLELPVKYTVYTLISROQDSTNH-VNFSSS---HGRROEAAHRYRVNLS 959  
 Qy 998 FPIHGMWKITPIATRSQNLRLKLRDLFTDANTSCNIWGNSTERYPTVEE-----DL 1052  
 Db 960 LKL-AVRNFWFVPL-----LNGVAVDWITLSSPAQVSCVSKMPQNPDLTQI 1009  
 Qy 1053 RRAPQLNHSDDVYSINCLRLVPNQ-EINPHLLGNL---WLRLS---KALKYKSMKIMV 1105  
 Db 1010 QRRSVLDCSTADCLHFRCDIPSLDIQDELDFILGNLSFGWVSQTLQEKVLLVSEAITP 1069  
 Qy 1106 NAALQRFHSPFIREEDPSQIEFEISKOEDWV---PIMLVGSTLGGILLALLALVAL 1163  
 Db 1070 DTSVSLQPGQEAFLR---AQVETTL---EYVVYEPFLVAGSSVGGILLALLATVVL 1122  
 Qy 1164 RKLGFTRSAARRRPPGLD 1181  
 Db 1123 YKLG---XKQYKEMLD 1137

RESULT 13  
 US-09-350-259-55  
 ; Sequence 55, Application US/09350259  
 ; Patent No. 6620915  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gallatin, Michael W.  
 ; TITLE OF INVENTION: No. 8620915el Human 2  
 ; FILE REFERENCE: 27866/35004  
 ; CURRENT APPLICATION NUMBER: US/09/350,259  
 ; CURRENT FILING DATE: 1999-07-08  
 ; EARLIER APPLICATION NUMBER: 09/193,043  
 ; EARLIER FILING DATE: 1998-11-16  
 ; EARLIER APPLICATION NUMBER: 08/173,497  
 ; EARLIER FILING DATE: 1993-12-23  
 ; EARLIER APPLICATION NUMBER: 08/286,889  
 ; EARLIER FILING DATE: 1994-08-05  
 ; EARLIER APPLICATION NUMBER: 08/362,652  
 ; EARLIER FILING DATE: 1994-12-21  
 ; EARLIER APPLICATION NUMBER: 08/943,363  
 ; EARLIER FILING DATE: 1997-10-03  
 ; NUMBER OF SEQ ID NOS: 114  
 ; SOFTWARE: Patent in Ver. 2.0  
 ; SEQ ID NO 55  
 ; LENGTH: 1161  
 ; TYPE: PRT  
 ; ORGANISM: Rattus rattus  
 US-09-350-259-55

Query Match 19.0%; Score 1184.5; DB 4; Length 1161;

Best Local Similarity 29.6%; Pred. No. 1,8e-97;  
 Matches 361; Conservative 207; Mismatches 519; Indels 131; Gaps 43;  
 Qy 7 LVVAWALSMPGFTDTFNDRKPRVPGSTAFPGYTVQOHDISGNKWLVTGALPTNG 66  
 Db 8 LLCGWLASCHG-----SNLDVEEP-IVFREDAAFGQTVQF---GGSELVVGAFLA 59  
 Qy 67 YQKTGDVYKCPVIHGNCTKLNALGRVTLNSVSRKKNMRLGSLATNPKNSFLACSPWS 126  
 Db 60 VNQTRLVDCAPATWCQPIVL-RSPLEAV-----NMSGLSLVATATNNAQLLACGPTAQ 113  
 Qy 127 HECGSSYYTTGCMGRVNSNFRSKTVAPALQRC-QTYMDIVIVLDSGNSI---YPWVEVOH 183  
 Db 114 RACVKNMYAKGSCLLGSSLSQFIQAVPASMPCEPRQEMDIAFLIDGSGINORDFAQMKD 173  
 Qy 184 FLINILAKFYIGPGQIQGVVQYQGEDVHEPHLNDYRSKVVVBAASHIEQRGGTETRTA 243  
 Db 174 FYKALMGF--ASTSTLSLMQISNLKTHFTTFEKNILDPQSLVDPIVQLQGL-TYTA 230  
 Qy 244 FGIEFARSEAF--QXGRRGAKKVMVITDGSNDSP-DLEKVIQOESRDNTRVAVAVL 300  
 Db 231 TGIPTVMEELFHSKNGSRKSKILLVITDQKRDPLEYSDVIPAADKAGIRYAVGV 290  
 Qy 301 GYNNRGINPETFINEIKVILASDDPKHFNVTDEAALKDIVDALGDRIFSLSGT-NKXE 359  
 Db 291 DAFQE-----PTALKELNTIGSAPPQDHRVKNFAALRSIORQOEKIFALEGTQSRSS 345  
 Qy 360 TSFGLMSQTSFSSHVVEDGVLGAVGAYDNWAGVAKTSAGKVIP--LRSEYLKPEFPEE 417  
 Db 346 SSGFQHEMSQEGSSALTSQDGLVAGVSGFSWGGAF-----LYPPTTRPTFNMNQEN 398  
 Qy 418 LKHGAYLYGTYTTSVSSQGRVYVAGAPRPHNTKVLFTWHNRESLTIHOAMRQOIG 477  
 Db 399 VDMRDSYLGYS-TAVAFMKGVHSLIIGAPRHQHTKQWIFT-QEAKHWKPKSEVRQTQIG 456  
 Qy 478 SYFGSEITSDIDGQVTVLLVGAPEYFNEGRKGVVYVELR--QNEFVYNGTLDKSH 535  
 Db 457 SYFGSLCSVDVDRGSDXDLVIGAPHYYEQTR-GQVSVFPPVGVGRWQCEATLHGEQ 515  
 Qy 536 SYQNAAPGSSIASVRDLNODSYNDVVVGAPELNDHAGAIYIFHG-FRGSILKTPKRIIA 594  
 Db 516 GHFWGRFGVALTVLGDVNGDNLADVAIGAPEGSEERGAIVYIFEGASRLRIMPSPSRVTG 575  
 Qy 595 SELATCLOYGSGSIHQDLNEDGLDLAVGALGNVILWSRPPVAINASLPEPSKINI 654  
 Db 576 SLSLRLQYFGQSLGGQDITQDGLVLAQAGHVLRLSLPLKVKELSIKPAPEVAK 635  
 Qy 655 FHRDCRSGRDATCL---AAFLCFTPIFLAPHFQTTTGV-IRYNATMDERRYPRAHLDE 710  
 Db 636 AVQCWE--RTPTVLEAGEATVCLTVHKSGPDLGNGVQSVRYDLALDPCRLISRAIFDE 693  
 Qy 711 GDRFTNRAVLSSGOELCERINPHVLD-TADYVKPVTFSVEYSL-----EDPDHGMPLD 764  
 Db 694 TKNCTLTGRKTLGLGDH-CETVKLLLPDCEVEDAVSPIILRLNLSLRVDSASPNLHPVLA 752  
 Qy 765 DGNPTTLRVSPVFWNGCNDEHCVDPDLVDARSDELPTAMEYCORVLRKPAQDCSAYTLSP 824  
 Db 753 VGSQDHIATSLPFKNCKQELCEGDL-----GISFNF 785  
 Qy 825 DTTVTIESTRQVAVATLENGENAYSTVLNISQSANLQFASLI---QKSDSGSIEC 881  
 Db 786 SGLQVLVVGSPSLTIVTVWNEGEDSYGLTVKFPYFAGLSYRVGTQOHPYPLRLAC 845  
 Qy 882 VNE---ERRLQKQVNCVSYFFRAKAKAVAFRLDSEFSKIFL-HHLEIELAAGSDNERD 937  
 Db 846 EAEPAQEDLRSSSCSINHPIFREGAKTTFWITFDVSKAFGLDRLLRKAKASSENKPD 905  
 Qy 938 STKEDNVAPLRFLKYRADVLFTRSSLSHYEVLKNSLERYDGIOPPFSCIFRIONLGL 997  
 Db 906 TNK--TAFQLELPVKYTVYTLISROQDSTNH-VNFSSS---HGRROEAAHRYRVNLS 959  
 Qy 998 FPIHGMWKITPIATRSQNLRLKLRDLFTDANTSCNIWGNSTERYPTVEE-----DL 1052

Db 960 LKL-AVRNFWPVL-----LNGVAVDVTLSSPAQGVSVSQMKPQNDPLTQI 1009  
QY 1053 RRAPOLNHSNDSVINCINRILVPNQ-EINEHLGNL---WLSL---KALKYKSMKMW 1105  
Db 1010 QRRSLDCSIADCLFRCDIPSLD:QBELDFILRGKLSFGVWSQTLQEKVLLVSEAITF 1069  
QY 1106 NAALQORQSPRIFREEDPSRQIEPEISKQEDMOV--PIWIVGSTLGLLLALLLVAL 1163  
Db 1070 DTSVYSQLEQBAFLR---AQVEITL---BEYVVEPIFLVAGSVGGLLLALLITVL 1122  
QY 1164 KLGFRSARRRREGLD 1181  
Db 1123 YKLGf---XKROYKEMLD 1137

## RESULT 14

US-08-286-889-46  
; Sequence 46, Application US/08286889  
; Patent No. 5470953  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Mich  
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: P38,659  
; REFERENCE/DOCKET NUMBER: 27866/32168  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1155 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-286-889-46

Query Match 19.0%; Score 1181.5; DB 1; Length 1155;  
Best Local Similarity 30.0%; Pred. No. 3.4e-97;  
Matches 373; Conservative 199; Mismatches 501; Indels 171; Gaps 50;  
QY 5 RGLV---AWALSRLWPGFTDFTMDTRKPRVPGSRTAFPGYTVQHQHDSGNKMLVVGAP 61  
Db 3 RGVVILLCWALASCHG-----SNLDVEKP-VVFKEDASFGTVVQF---GGSRLVVGAP 54  
QY 62 LFTNGYQKTDGVYKCPVINGNCTKLNLRVTLSNVSERKDNRLGLSLATNPKNISFLAC 121  
Db 55 LEAVAVNQTGSSDCPPARGVCQPIILL-HIPLEAV-----NMSGLSLVADTNNSQLAC 108  
QY 122 SPLWSEHCSSYTTGMCNRSVNSNFRFSKTVAPALQRCQ-TYMDIVIVLDGNSNI--YPW 178

Db 109 GPTAQRACAKMYAKGSCLLLGSSLOFIQAIPATMPECPCQEMDIAFLDGSISDOSDF 168  
QY 179 VEVQHFLINILKXFIYGPQI-----QGVVVQYGEDVVHFLNDYRSKVDVVAASHIE 233  
Db 169 TOMKDFVKALM-----GOLASTSTSFSLAQSYNLIKTHFTPTTFKSSLSQSLVAIV 221  
QY 234 QRGTETRTAFGIEFARSEAF--CKGGRKGAKKMWIVITDGHSDSP-DLEKVIQOQSRD 290  
Db 222 QLOQL-TYTAGGKQKVKELFHSKNGARKSAKILIVITDQKFRDPLEYHVLPEAKA 280  
QY 291 NVTRYAVAVLGYNNRGINPETFLNEIKYIASOPDDKHFFNVTDAAALKDIDVALGDRIF 350  
Db 281 GIIRYAIAGVGAFAE-----PTALQELNTGSSAPSDQHVFKVGNFVALRSIORIQEKIF 335  
QY 351 SLEGT-NKNETSFGLMSQTSFSSHVVVEDGVLLGAVGAYDNGAVLAKETSAGKVIPLRES 409  
Db 336 AIEGTSSSSSSPOHEMSQEGFSALSNDGFLVAGVGSWGAFLYPS-----NKRST 390  
QY 410 YLKEFFPEELKNHGAYLGYTVTSVSSRQGVYVAGAPRPNHTKVLFTMNNRSLTHQ 469  
Db 391 FINMSQENEDMRDAYLGS-TALAFWKGVHSLILGAPRQHTGKVIPT-QESRHRPKS 448  
QY 470 AMRQOIGSYFGSEITSDVIDGQVTDVLLVGAEMYNEGRERKGVYVELR--ONRVY 527  
Db 449 EVRGTOIGSYFGASLCSVDMDRDGSITDLVIGVPHYETHR-GGQSVCPMPGVRSRWHC 507  
QY 528 NGTLKDSHSYQNAFSGSSIASVRDLNQDSYNDVVVVGAPLEBNDHAGAIYIPHG-FRGSILK 586  
Db 508 GTTLHGEGHPWGRFGAALTVLGDVNGDSLADVAIGAPEGREENRGAVYIPHGASRQDLAP 567  
QY 587 TPQORITASELATGLOYEGCSIHQOLDNEDGLDIAVAGLGNNAVILSRVVOINSLH 646  
Db 568 SPSQRTVGSQFLRLQYFGQSLSGQDLTQGLDVLAVAGQGHVLLRSLPLKVGISIR 627  
QY 647 FEPSKINIFHRDCKRSGRDATCL---AAFCLFT-----PIPLAPHFQTTTIGIRYNAMD 698  
Db 628 FAPSEVAKTVYQC--WGRTPTVLEAGEATVCLTVRKSGSPDLG---DVQSSVAYDLALD 681  
QY 699 ERYTPRAHLDEGGDRPTNRAVLLSSGOELCERINFHVL-DAYVVPVTSVEYSL---754  
Db 682 PGRILISRAIFDETRKNTLTKRKTGLGDH-CETHKLLLPDCVEDAVTPIILRLSLAGD 740  
QY 755 EDPDHG--PMLDDGWPTTLRVSVFPMGNCNEDECPDVLVDARSDLPTAMEYCORVLK 812  
Db 741 SAPSENLRPLVAVGSDHVTASPPF-----KXNCEGNL-----773  
QY 813 PAQDCSAYTLGSDTTVFIIBSTRQRAVEATLENRGENAYSTVLNISQSANLQFASLIQK 872  
Db 774 -----GVSNFSGQLQVLEVGSSDELATVTVTWNEGEDSYGLIKFYYPABLSYRVTRA 827  
QY 873 ESDSG---SIECVNE---ERRLOQVCNVSYPFERAKAKVAFRLDSSEKSIPL-HLEI 925  
Db 828 QQPHYPILRLACEABPTGQESLRSSCSINHPIREGAKATFMITFDVSKAFGLDRLLL 887  
QY 926 ELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFRSS-----SLSHYEVKLNLSLE 977  
Db 888 RASASSENKPKTSK--TAFQLELPVKYTVTVVISRQEDSTKTFNFSSSHGE-RQKZAEH 944  
QY 978 RYDGIQPPFCIFRIQNLGLFPFHGMKMTIPTATRSQNL--LKLRDPLTDEANTSCN 1035  
Db 945 RY-----RVNLSPLTL-AISVNFVWPILL-NGVAVMDVTLR---SPAQGVSC-987  
QY 1036 IWNSTEVPTFVREEDURRAPO-----LNHSNSDVVSINCNILRVLPQNEINFHLGNL--1088  
Db 988 -----VQRBPQHSDDLTLTOIGESVLDCAIDCLHLRCDIPSLGTLDELDFILKGNLSF 1042  
QY 1089 WLSL---KALKYKSMKIMYNAALQRFHSPTIFREEDPSRQIEPEISKQEDMOVVI 1144  
Db 1043 GWSIQTLQKVKLLSEAEITFNTSVYSQLPQGEAFRAQVSTMLEEYVVE-----PVFL 1097  
QY 1145 IVGSTLGLLLALLLVALRKLGFPSARRERE-----PGLDPTP 1184  
Db 1098 WPFSSVGGLELLALITVALYKLGFFK--RQYKEMLDLPSADPDP 1139

## RESULT 15

US-08-485-618-46  
 ; Sequence 46, Application US/08485618  
 ; Patent No. 5728533  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gallatin, W. Michael  
 ; APPLICANT: Van der Vlieren, Monica  
 ; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit  
 ; NUMBER OF SEQUENCES: 103  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 ; STREET: 233 South Wacker Drive, 6300 Sear Tower  
 ; City: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: United States  
 ; ZIP: 60606-6402  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/485,618  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/173,497  
 ; FILING DATE: 23-DEC-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/286,889  
 ; FILING DATE: 5-AUG-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/362,652  
 ; FILING DATE: 21-DEC-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Williams Jr., Joseph A.  
 ; REGISTRATION NUMBER: 38,659  
 ; REFERENCE/DOCKET NUMBER: 27866/32797  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312-474-6300  
 ; TELEFAX: 312-474-0448  
 ; TELEX: 25-3856  
 ; INFORMATION FOR SEQ ID NO: 46:  
 ; TYPE: amino acid  
 ; LENGTH: 1155 amino acids  
 ; SEQUENCE CHARACTERISTICS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-485-618-46

Query Match 19.0%; Score 1181.5; DB 1; Length 1155;  
 Best Local Similarity 30.0%; Pred. No. 3.4e-97;  
 Matches 373; Conservative 199; Mismatches 501; Indels 171; Gaps 50;

QY 5 RGLV---AWALSLWPGFTDFNMTRKPRVPGSRTPAFGTVVQOHDISGNKWLWVGAP 61  
 DB 3 RGVVILLCGWALASCHG---SNLDVEKP-VVFKEDAASFGQTVVQF---GGRLVVGAP 54  
 QY 62 LEFNGYQKTDVYKCPVIHGNCTKMLGRTVLSNVSERKONRMLGLSLATNPDKNSFLAC 121  
 DB 55 LEAVAVNQSGSDCPATGVQCPIILL-HIPLEAV-----NMSLGLSLVADTNNSQLLAC 108  
 QY 122 SPLWSECGSSYYTTCMCGRVSNRPNRFSKTVAPALQRCQ-TYMDIVIVLDGNSI--YPW 178  
 DB 109 GPTAQACAKNMYAKSGCLLGSSLQFIQAIPTMPCPCPGQEMDIAFLIDGSGSIDQSD 168  
 QY 179 VEVOHFLINILKKFYTGPOI-----QGVWQGEDVVEHFNLDYRSVQVWVEAASHIE 233  
 DB 169 TQKDFVKALM-----GQLASTSTSFSLMQYSNLIKTHFTTFEKSLSQSLVDAIV 221  
 QY 234 QRGGTETRTAFGIEFARSEAF--QKGRGKGAKMIVITDGSBHDSP-DLEKVIQDSERD 290

Search completed: June 24, 2004, 17:56:46  
 Job time : 28 secs

Db 222 QLQGL-TYTAGIQKVKELFHSKKGAKSAKKLIVITDGGKFDPLRYRVIPEAKA 280  
 QY 291 NVTRYAVAVLGYNNRGINPTFFLNEIKYIASDPDOKHFFNVTDAAKIDVADLGRIF 350  
 Db 281 GIIRYAIGVGDAFRE-----PTALQELNTIGSAPSQSDHVFVKGNFVALRSIQRIQEKIF 335  
 QY 351 SLEGT-NKNETSPGLEMSQCTGPFSSHVVEDGVLLGAVGAYDMNGAVLKETSACKVPIRES 409  
 Db 336 ALEGTESSSSSFQHEMSQEGFSSALSMDGPVLGAVGGFSWGGAFILPS-----NMNST 390  
 QY 410 YLKEPPEELKNHGYLYGTYTTSVSSRQGRVYVAGAPRNFHTGKVLTFMHNRLTHQ 469  
 Db 391 FINMSQENEDMDADYLGYS-TALAFWKGVHSLILGAPRHOHTGKVIFT-QESRWRPKS 448  
 QY 470 AMRGOIGSYRGSEITSDVDDGVTVDLLACAPYFNEGRGKVVYELR--QNRFY 527  
 Db 449 EVRGTIQISYFGSLCSVDMDRGDGTDLVLGVPHYEHTR-GGVSVCPMPGVSRWHC 507  
 QY 528 NGTLKSHSYONARFGSSIASVRDINODSYNDVVVVGAPLEDNHAGAIYIFHG-FRGSILK 586  
 Db 508 GTTLHGEQHPWRFGAALTVLGVDNGDSLADVAIGAPGEENRGAVYIFHGASQDIAP 567  
 QY 587 TPQGRITASELATGLOYECCSIHQGLDNEGHLDLAVGALGNVILHSRPVQVQINASH 646  
 Db 568 SPQQRVTGSLFLRLQYFGQSLGGQDLTQDGLVDLAVGAGQGVLLSLPLKVGISIR 627  
 QY 647 FEPSKINIFHRDCKSGSDATCL---AAFLCFT-----PIFLAPHFOTTTVGIRYNAMD 698  
 Db 628 FAPSEVAKTVQC--WGRTPTVLEAGEATVCLTVKRGSPDLG---DVQSVRYDLALD 681  
 QY 699 ERYTPRAHLDGGDRFTNRAVILLSSGOELCERINFHVLDTADYVKVTVTSVEVSL--- 754  
 Db 682 FORLISRAIFDETKNCTLTTRKTLGLGDH-CETMKLLLPDCVEDAVTPIILRLNLSAGD 740  
 QY 755 EDPDHG--PMLDDGWPTTLRVSVFPWNGCNDHCVDPDLVDLSDARSDLPTAMEYCORVLR 812  
 Db 741 SAPSRNLAPVLAVGSDHVTASFPF-----EKXCEGNL----- 773  
 QY 813 PAQDCSAYTLSPDITVFIESTRORVAVEATLENRGENAYSTVLNISANLQFASLQK 872  
 Db 774 -----GVSENFSGQLVLEVSGSPELTVTVTMNEGEDSYGLTKFYYPAELSYRRVTRA 827  
 QY 873 EDSDG---SIECVNB---ERLQKQVCNVSPFFRAKAKVAFRLDSEPSKSIPL-HHLEI 925  
 Db 828 QQPHYPPLRLACEAPTEQESLRSSSCSINHPIFRGAKATPMITFDVSYKAFGLDRLLL 887  
 QY 926 ELAAGSDSNRSTKEDNVAPLRFHLKYEADVLFTRSS-----SLSHYEVKXNSLSLE 977  
 Db 888 RASASSENKPKTSK--TAFQLELPVKYTVTVTVISROEDSTKHFNFSHSGE-ROKEAEH 944  
 QY 978 RYDGIQPPFSCIFRIQNLGLFPPIHGMMKTIPIATRSGNEL--LKLRDLTDEANTSCN 1035  
 Db 945 RY-----RVNLSPLTL-AISVNFVPIILL-NGVAVMDVTLR---SPAQGVSC- 987  
 QY 1036 IWGNSTYRPTTVBEDLERAPQ-----LNHNSDVVSINCNIT-RLVPQOEINPHLLGNL-- 1089  
 Db 988 -----VSQREPPQSHDLTLQIGRSVLDCAIDCLHLRCDIPSLGTLDLDELFLKGNLSF 1042  
 QY 1089 -WLASL--KAIKYKSMKIMVNAALQRFHSPFPIFREDDPSQIEFEISKQEDQVPTMI 1144  
 Db 1043 GWISQTLQKXVLLSEAEITNTSVISQLPQOEAFRAQVSTMLSEYVYV--PVFL 1097  
 QY 1145 IVGSTLGLLLALLVLAIRLKLGFRRARRR-----PGLDPTP 1184  
 Db 1098 MVSFSSVGLLLALLITVALYKLGFFK--RQVKEMLDLPSADPDP 1139

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 17:58:06 ; Search time 27 Seconds  
(without alignments)  
4232.427 Million cell updates/sec

Title: US-09-980-403-2  
Perfect score: 1188  
Sequence: 1 MDLPRGLVVAWLSLWPGFT.....FSSARRRRRGLDTPFKVLE 1188

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : PIR 78:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	1.3	272	2 A55348	integrin alpha-1 -
2	15	1.3	1151	2 A45226	integrin alpha-1 c
3	15	1.3	1180	2 A35854	integrin alpha-1 c
4	11	0.9	1170	2 I45914	integrin alpha-2 s
5	11	0.9	1181	2 A33998	integrin alpha-2 s
6	9	0.8	74	2 I51524	integrin alpha 2 s
7	9	0.8	315	2 A35567	permease (imported
8	9	0.8	371	1 T43407	3-isopropylmalate
9	9	0.8	607	2 S60658	legumin - Gnetum g
10	9	0.8	1178	2 S44142	VLA-2 protein homo
11	8	0.7	20	2 A60822	cytochrome P450 PB
12	8	0.7	42	2 A34259	cytochrome P450mt4
13	8	0.7	76	2 I51527	integrin alpha 5 s
14	8	0.7	76	2 A45337	heat-stable antige
15	8	0.7	76	2 I53107	CD24 precursor - r
16	8	0.7	80	2 A48956	8 cell surface ant
17	8	0.7	103	2 AF0856	conserved hypoteth
18	8	0.7	127	2 B75301	hypothetical prote
19	8	0.7	141	2 I51785	heat-stable antige
20	8	0.7	224	2 E71228	hypothetical prote
21	8	0.7	226	2 S76800	hypothetical prote
22	8	0.7	234	1 S15102	eosinophil major b
23	8	0.7	264	2 S22090	catechol O-methyl
24	8	0.7	308	2 AG2637	conserved hypoteth
25	8	0.7	356	2 P97419	BH1459 conserved h
26	8	0.7	360	2 A85016	hypothetical prote
27	8	0.7	367	2 S19172	cytochrome P450 2B
28	8	0.7	387	1 DYNH4	dopamine receptor
29	8	0.7	402	1 S23860	chloramphenicol re

30	8	0.7	415	2 G83569	probable permeal
31	8	0.7	442	2 AG3504	dihydrofolate synt
32	8	0.7	444	2 T01721	hypothetical prote
33	8	0.7	487	2 T47107	benzaldehyde dehyd
34	8	0.7	491	1 O4RTPB	cytochrome P450 2B
35	8	0.7	491	1 O4RTP2	cytochrome P450 2B
36	8	0.7	500	2 B31047	testosterone 16alp
37	8	0.7	547	2 A31314	maioalactic enzyme
38	8	0.7	547	2 A31686	maioalactic enzyme
39	8	0.7	605	1 W1WLB	E1 protein - bovin
40	8	0.7	606	2 A72429	oligopeptide ABC t
41	8	0.7	614	2 A69845	Na+/H+ antiporter
42	8	0.7	620	1 W1WLB2	E1 protein - bovin
43	8	0.7	685	2 AC0527	ferichrome transp
44	8	0.7	697	2 T16306	hypothetical prote
45	8	0.7	739	2 A88445	protein C36E6.10 l
46	8	0.7	849	2 B83349	probable C1pA/B-t
47	8	0.7	1041	2 T31437	integrin alpha cha
48	8	0.7	1065	2 T25068	hypothetical prote
49	8	0.7	1146	2 S40311	integrin - fruit f
50	8	0.7	1153	1 RWHUB	cell surface glyco
51	8	0.7	1163	1 RWHUB	cell surface glyco
52	8	0.7	1179	2 A53213	integrin alpha-B c
53	8	0.7	1366	2 S7664	IGA-specific metal
54	8	0.7	1460	2 D81675	polymorphic membra
55	8	0.7	1664	2 S67250	DNA-directed RNA p
56	8	0.7	1802	2 H88444	protein C26E6.12 l
57	8	0.7	2238	1 RRVUEY	genome polypotein
58	7	0.6	47	2 D81686	hypothetical prote
59	7	0.6	74	2 T47376	hypothetical prote
60	7	0.6	89	2 I49515	B144 protein B - m
61	7	0.6	94	2 A99774	probable excisiona
62	7	0.6	99	2 P85636	probable excisiona
63	7	0.6	103	2 G75257	hypothetical prote
64	7	0.6	104	2 PH0141	integrin alpha cha
65	7	0.6	106	4 S57386	hypothetical prote
66	7	0.6	107	2 JC1127	major allergen cha
67	7	0.6	109	2 C56413	insulin precursor
68	7	0.6	110	1 INRB	insulin precursor
69	7	0.6	110	2 S30268	protein hdea precu
70	7	0.6	110	2 F91177	protein hdea precu
71	7	0.6	110	2 G86023	protein hdea precu
72	7	0.6	110	2 G86023	hypothetical prote
73	7	0.6	114	2 C95870	hypothetical prote
74	7	0.6	115	2 D75434	conserved hypoteth
75	7	0.6	132	2 B81215	hypothetical prote
76	7	0.6	133	2 H75386	hypothetical prote
77	7	0.6	139	1 C45051	lamprin 2 precursor
78	7	0.6	139	2 A38612	insulin-like growt
79	7	0.6	141	1 HACH1	hemoglobin alpha-D
80	7	0.6	141	2 S56103	hemoglobin alpha-D
81	7	0.6	142	2 JT0573	retinoic acid-indu
82	7	0.6	142	2 G87647	hypothetical prote
83	7	0.6	143	2 T05564	hypothetical prote
84	7	0.6	154	2 S14947	2S albumin - Brazi
85	7	0.6	154	2 F83074	conserved hypoteth
86	7	0.6	158	2 G82494	conserved hypoteth
87	7	0.6	160	1 P70351	NADH2 dehydrogenas
88	7	0.6	160	2 S58214	apoptosis inducer
89	7	0.6	161	2 D71903	hit family protein
90	7	0.6	163	2 S66795	probable membrane
91	7	0.6	163	2 E72225	conserved hypoteth
92	7	0.6	171	1 B64498	conserved hypoteth
93	7	0.6	176	2 AC0207	probable exported
94	7	0.6	176	2 T38925	cytochrome oxidase
95	7	0.6	176	2 T46709	hypothetical prote
96	7	0.6	176	2 A61250	hypothetical prote
97	7	0.6	179	2 A64551	hypothetical prote
98	7	0.6	179	2 D71957	hypothetical prote
99	7	0.6	180	2 E70358	HUPE hydroxgenase r
100	7	0.6	182	2 E82367	hypothetical prote

## ALIGNMENTS

RESULT 1  
 A55348  
 integrin alpha-1 - chicken (fragment)  
 C:Species: Gallus gallus (chicken)  
 C>Date: 10-Feb-1995 #sequence\_revision 10-Feb-1995 #text\_change 15-Sep-2003  
 C:Accession: A55348  
 R:Kern, A.; Briesewitz, R.; Bank, I.; Marcantonio, E.E.  
 J. Biol. Chem. 269, 22811-22816, 1994  
 A>Title: The role of the I domain in ligand binding of the human integrin alpha-1beta-1.  
 A:Reference number: A55348; MUID:94357930; PMID:7521332  
 A:Accession: A55348  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-272 <KER>  
 A:Cross-references: GB:U0114  
 F:55-230/Domain: von Willebrand factor type A repeat homology <VWA2>  
 Query Match 1.3%; Score 15; DB 2; Length 272;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 164 DIVIVLDGNSIYPW 178  
 DB 57 DIVIVLDGNSIYPW 71

RESULT 2  
 A45226  
 integrin alpha-1 chain - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 15-Sep-2003  
 C:Accession: A45226  
 R:Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.  
 J. Biol. Chem. 268, 2989-2996, 1993  
 A>Title: Expression of native and truncated forms of the human integrin alpha 1 subunit.  
 A:Reference number: A45226; MUID:9315124; PMID:8428973  
 A:Accession: A45226  
 A>Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-1151 <BRI>  
 A:Experimental source: hepatoblastoma cell line HepG2  
 A>Note: sequence extracted from NCBI backbone (NCBI:P124326)  
 F:142-317/Domain: von Willebrand factor type A repeat homology <VWA1>  
 Query Match 1.3%; Score 15; DB 2; Length 1151;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 164 DIVIVLDGNSIYPW 178  
 DB 144 DIVIVLDGNSIYPW 158

RESULT 3  
 A35854  
 integrin alpha-1 chain precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 23-Oct-1990 #sequence\_revision 13-Sep-1991 #text\_change 15-Sep-2003  
 C:Accession: A35854; S11243  
 R:Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J. Cell Biol. 111, 709-720, 1990  
 A>Title: Molecular cloning of the rat integrin alpha-1 subunit: a receptor for laminin a  
 A:Reference number: A35854; MUID:90338125; PMID:2380249  
 A:Accession: A35854  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1180 <IGN>  
 A:Cross-references: GB:X52140; NID:G56493; PID:CAA36384.1; PID:G56494  
 C:Keywords: cell adhesion; cytoskeleton; transmembrane protein  
 F:170-345/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 1.3%; Score 15; DB 2; Length 1180;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 164 DIVIVLDGNSIYPW 178  
 DB 172 DIVIVLDGNSIYPW 186

RESULT 4  
 I45914  
 integrin alpha 2 subunit - bovine (fragment)  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 15-Sep-2003  
 C:Accession: I45914  
 R:Kamata, T.; Puzon, W.; Takada, Y.  
 J. Biol. Chem. 269, 9659-9663, 1994  
 A>Title: Identification of putative ligand binding sites within the I-domain of integ  
 A:Reference number: A54402; MUID:94193647; PMID:7511592  
 A:Accession: I45914  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1170 <KAM>  
 A:Cross-references: GB:L25886; NID:G439695; PID:AA59595.1; PID:G439696  
 F:161-336/Domain: von Willebrand factor type A repeat homology <VWA2>  
 Query Match 0.9%; Score 11; DB 2; Length 1170;  
 Best Local Similarity 100.0%; Pred. No. 0.071;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAPMY 505  
 DB 495 TDVLLVGAPMY 505

RESULT 5  
 A33998  
 integrin alpha-2 chain precursor - human  
 N:Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Mar-1990 #sequence\_revision 18-Sep-1992 #text\_change 15-Sep-2003  
 C:Accession: A33998; B56793; A53117  
 R:Takada, Y.; Hemler, M.E.  
 J. Cell Biol. 109, 397-407, 1989  
 A>Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (plate  
 A:Reference number: A33998; MUID:89308879; PMID:2545729  
 A:Accession: A33998  
 A:Molecule type: mRNA  
 A:Residues: 1-1181 <TAK>  
 A:Cross-references: GB:X17033; NID:G33906; PID:CAA34894.1; PID:G33907  
 A>Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue  
 R:Catimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.  
 Biochem. J. 279, 419-425, 1991  
 A>Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIC\*, GPIIa  
 A:Reference number: A56793; MUID:92061944; PMID:1953640  
 A:Accession: B56793  
 A:Molecule type: protein  
 A:Residues: 30-43 <CAT>  
 A:Experimental source: platelet  
 R:Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Teung, Y.L.; Gafford, A.  
 J. Biol. Chem. 269, 463-469, 1994  
 A>Title: The human alpha-2 integrin gene promoter. Identification of positive and neg  
 A:Reference number: A53117; MUID:94103255; PMID:8276836  
 A:Accession: A53117  
 A:Molecule type: DNA  
 A:Residues: 1-16, 'V', 18-21 <ZUT>  
 A:Cross-references: GB:L24121; NID:G400342; PID:AA16519.2; PID:G4583535  
 A>Note: authors translated the codon GTA for residue 17 as Leu  
 C:Genetics:  
 A:Gene: GDB:ITGA2; CD49B  
 A:Cross-references: GDB:128031; OMIM:192974  
 A:Map position: 5q11.1-5q11.2

C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein  
 F;1-29/Domain: signal sequence #status predicted <SIG>  
 F;30-1133/Domain: extracellular #status predicted <EXT>  
 F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>  
 F;1134-1154/Domain: transmembrane #status predicted <TM>  
 F;1155-1181/Domain: intracellular #status predicted <CYT>  
 F;1105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 0.9%; Score 11; DB 2; Length 1191;  
 Best Local Similarity 100.0%; Pred. No. 0.072;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 495 TDVLLVGAPMY 505  
 |||||  
 Db 506 TDVLLVGAPMY 516

## RESULT 6

I51524  
 integrin alpha 2 subunit - African clawed frog (fragment)  
 C;Species: Xenopus laevis (African clawed frog)  
 C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 15-Sep-2003  
 C;Accession: I51524  
 R;Whittaker, C.A.; DeSimone, D.W.  
 Development 117, 1239-1249, 1993  
 A;Title: Integrin alpha subunit mRNAs are differentially expressed in early Xenopus embryos  
 A;Reference number: I51524; MUID:94008528; PMID:8404528  
 A;Accession: I51524  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-74 <WHI>  
 A;Cross-references: GB:I10186; NID:g214538; PIDN:AAA16246.1; PID:g214539

Query Match 0.8%; Score 9; DB 2; Length 74;  
 Best Local Similarity 100.0%; Pred. No. 0.67;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 495 TDVLLVGAP 503  
 |||||  
 Db 14 TDVLLVGAP 22

## RESULT 7

AH3567  
 permease [imported] - Brucella melitensis (strain 16M)  
 C;Species: Brucella melitensis  
 C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
 R;DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, C.; Los, T.; Ivanova, .; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
 A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
 A;Reference number: AD3252; PMID:11756688  
 A;Accession: AH3567  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-315 <NUR>  
 A;Cross-references: GB:AB008918; PIDN:AAL53707.1; PID:g17984630; GSPDB:GN00191  
 A;Experimental source: strain 16M  
 C;Genetics:  
 A;Gene: BME110465  
 A;Map position: 11

Query Match 0.8%; Score 9; DB 2; Length 315;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLALLVLA 1162  
 |||||  
 Db 283 LLALLVLA 291

## RESULT 8

Query Match 0.8%; Score 9; DB 2; Length 607;  
 Best Local Similarity 100.0%; Pred. No. 4.2;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

T43407  
 3-isopropylmalate dehydrogenase [EC 1.1.1.85] [validated] - fission yeast (Schizosaccharomyces pombe)  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 21-Jul-2000  
 C;Accession: T43407; T39850; T39871  
 R;Kikuchi, Y.; Kitazawa, Y.; Shimatake, G.; Yamamoto, M.  
 Curr. Genet. 14, 375-379, 1988  
 A;Title: The primary structure of the leu1+ gene of Schizosaccharomyces pombe.  
 A;Reference number: Z22491; MUID:89106270; PMID:3063400  
 A;Accession: T43407  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-371 <KIK>  
 A;Cross-references: EMBL:M36910; NID:g173411; PIDN:AAA35316.1; PID:g173412  
 R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.  
 submitted to the EMBL Data Library, July 1998  
 A;Reference number: Z21885  
 A;Accession: T39850  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-371 <WOO>  
 A;Cross-references: EMBL:AL031174; PIDN:CAA20106.1; GSPDB:GN00067  
 A;Experimental source: strain 972h-; cosmid clA4  
 R;Volckaert, G.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, February 1998  
 A;Reference number: Z21887  
 A;Accession: T39871  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 152-371 <VOL>  
 A;Cross-references: EMBL:AL021746; PIDN:CAA16840.1; GSPDB:GN00067; SPDB:SPBCL1E9.07C  
 A;Experimental source: strain 972h-; cosmid clE8  
 C;Genetics:  
 A;Gene: leu1; SPDB:SPBCL1E8.07C; SPBCL1A4.02C  
 A;Map position: 2  
 C;Function: <IPD1>  
 A;Description: catalyzes the oxidation of 3-carboxy-2-hydroxy-4-methylpentanoate to 3  
 A;Pathway: leucine biosynthesis  
 C;Function: <IPD2>  
 A;Description: EC 1.1.1.85 [validated, MUID:89106270]  
 C;Superfamily: 3-isopropylmalate dehydrogenase  
 C;Keywords: oxidoreductase

Query Match 0.8%; Score 9; DB 1; Length 371;  
 Best Local Similarity 100.0%; Pred. No. 2.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 378 DGVLLGAVG 386  
 |||||  
 Db 69 DGVLLGAVG 77

## RESULT 9

S60658  
 legumin - Gnetum gnemon  
 N;Alternate names: 11S globulin; seed storage protein  
 C;Species: Gnetum gnemon  
 C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Aug-1999  
 C;Accession: S60658  
 R;Braun, H.; Horstmann, C.; Baumlein, H.  
 submitted to the EMBL Data Library, August 1995  
 A;Description: Legumins of the Gnetatae: characterization and evolutionary relationships  
 A;Reference number: S60658  
 A;Accession: S60658  
 A;Molecule type: mRNA  
 A;Residues: 1-607 <BRA>  
 A;Cross-references: EMBL:Z50779; NID:G949870; PIDN:CAA90642.1; PID:G949871  
 C;Superfamily: glycinin  
 C;Keywords: seed; storage protein

Query Match 0.8%; Score 9; DB 2; Length 607;  
 Best Local Similarity 100.0%; Pred. No. 4.2;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1161  
 |||||  
 Db 10 LLLALLLV 18

## RESULT 10

S44142  
 VLA-2 protein homolog - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 15-Sep-2003  
 C:Accession: S44142  
 R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze  
 submitted to the EMBL Data Library, January 1994  
 A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not  
 A:Reference number: S44142  
 A:Accession: S44142  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1178 <EDE>  
 A:Cross-references: EMBL:Z29987; NID:G473098; PIDN:CAA82877.1; PID:G473099  
 F:169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 0.8%; Score 9; DB 2; Length 1178;  
 Best Local Similarity 100.0%; Pred. No. 7.4; Mismatches 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0

QY 495 TDVLLVGAP 503  
 |||||  
 Db 503 TDVLLVGAP 511

## RESULT 11

A60822  
 cytochrome P450 PB-3a - rat (fragment)  
 N:Contains: oxidoreductase (EC 1.-.-.-)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 11-Jun-1999  
 C:Accession: A60822; I55191  
 R:Ameliazad, Z.; Narbonne, J.F.; Wolf, C.R.; Robertson, L.W.; Oesch, F.  
 Biochem. Pharmacol. 37, 3245-3249, 1988  
 A:Title: Effect of nutritional imbalances on cytochrome P-450 isozymes in rat liver.  
 A:Reference number: A60822; MUID:88293549; PMID:3041969  
 A:Accession: A60822  
 A:Status: preliminary  
 A:Molecule type: protein

A:Residues: 1-20 <AME>  
 R:Hashimoto, T.; Matsumoto, T.; Nishizawa, M.; Kawabata, S.  
 J. Biochem. 103, 487-492, 1988  
 A:Title: A mutant rat strain deficient in induction of a phenobarbital-inducible form of

A:Reference number: I55191; MUID:88273074; PMID:2839467  
 A:Accession: I55191  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-20 <RES>  
 A:Cross-references: GB:D00250; NID:G220727; PIDN:BA00181.1; PID:G220728  
 C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
 C:Keywords: electron transfer; heme; monooxygenase; oxidoreductase; transmembrane protein

Query Match 0.7%; Score 8; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 2.2; Mismatches 0; Indels 0; Gaps 0;  
 Matches 8; Conservative 0

QY 1153 LLLALLLV 1160  
 |||||  
 Db 6 LLLALLLV 13

## RESULT 12

A34259  
 cytochrome P450mt4, phenobarbital-inducible, mitochondrial, hepatic - rat (fragment)  
 N:Contains: oxidoreductase (EC 1.-.-.-)  
 C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 05-Mar-1999  
 C:Accession: A34259  
 R:Shayiq, R.M.; Avadhani, N.G.  
 Biochemistry 29, 866-873, 1990  
 A:Title: A phenobarbital-inducible hepatic mitochondrial cytochrome P-450 immunochemi  
 A:Reference number: A34259; MUID:90254127; PMID:2340279  
 A:Accession: A34259  
 A:Molecule type: protein  
 A:Residues: 1-42 <SHA>  
 C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
 C:Keywords: electron transfer; heme; liver; mitochondrion; monooxygenase; oxidoreduct

Query Match 0.7%; Score 8; DB 2; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 4.1; Mismatches 0; Indels 0; Gaps 0;  
 Matches 8; Conservative 0

QY 1153 LLLALLLV 1160  
 |||||  
 Db 6 LLLALLLV 13

## RESULT 13

I51527  
 integrin alpha 5 subunit - African clawed frog (fragment)  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 29-Sep-1999  
 C:Accession: I51527  
 R:Whittaker, C.A.; DeSimone, D.W.  
 Development 117, 1239-1249, 1993  
 A:Title: Integrin alpha subunit mRNAs are differentially expressed in early Xenopus e  
 A:Reference number: I51524; MUID:94008528; PMID:8404528  
 A:Accession: I51527  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-76 <WHI>  
 A:Cross-references: GB:L10191; NID:G214544; PIDN:AAA16249.1; PID:G214545  
 C:Superfamily: integrin alpha-2b chain

Query Match 0.7%; Score 8; DB 2; Length 76;  
 Best Local Similarity 100.0%; Pred. No. 6.9; Mismatches 0; Indels 0; Gaps 0;  
 Matches 8; Conservative 0

QY 541 RFGSSIAS 548  
 |||||  
 Db 63 RFGSSIAS 70

## RESULT 14

A43537  
 heat-stable antigen M1/69-J1ld precursor - mouse  
 N:Alternate names: CD24 protein; nectadrin  
 C:Species: Mus musculus (house mouse)  
 C:Date: 06-Nov-1992 #sequence\_revision 06-Nov-1992 #text\_change 31-Jan-2000  
 C:Accession: A43537; I48287; S15784; S15783; S43709; S32240; S33129  
 R:Kay, R.; Takeda, F.; Humphries, R.K.  
 J. Immunol. 145, 1952-1959, 1990  
 A:Title: Expression cloning of a cDNA encoding M1/69-J1ld heat-stable antigens.  
 A:Reference number: A43537; MUID:90361906; PMID:2118158  
 A:Accession: A43537  
 A:Molecule type: mRNA  
 A:Residues: 1-76 <KAY>  
 A:Cross-references: GB:M58661; NID:G198985; PIDN:AAA39481.1; PID:G198986  
 R:Wenger, R.H.; Rochelle, J.M.; Seidin, M.F.; Kohler, G.; Nielsen, P.J.  
 J. Biol. Chem. 268, 23345-23352, 1993  
 A:Title: The heat stable antigen (mouse CD24) gene is differentially regulated but ha  
 A:Reference number: A48876; MUID:94043127; PMID:8226859  
 A:Accession: I48287  
 A:Status: translation not shown; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-76 <RES>  
 A:Cross-references: EMBL:X79910; NID:G296466; PIDN:CAA51415.1; PID:G296467  
 R:Wenger, R.H.; Ayane, M.; Bose, R.; Koehler, G.; Nielsen, P.J.  
 Eur. J. Immunol. 21, 1039-1046, 1991

A:Title: The genes for a mouse hematopoietic differentiation marker called the heat-stab  
A:Reference number: S15783; MUID:91209380; PMID:2019286  
A:Accession: S15784  
A:Molecule type: DNA  
A:Residues: 1-76 <WE3>  
A:Cross-references: EMBL:X56469; NID:G51439; PIDN:CAA39841.1; PID:G51440  
A:Accession: S15783  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 32-76 <WE2>  
A:Cross-references: EMBL:X53825  
R.Nielsen, P.J.  
submitted to the EMBL Data Library, July 1990  
A:Reference number: S19111  
A:Accession: S43709  
A:Molecule type: mRNA  
A:Residues: 1-76 <WE>  
A:Cross-references: EMBL:X53825; NID:G55441; PIDN:CAA37822.1; PID:G55442  
C:Genetics:  
A:Gene: CD24a  
A:Map position: 10  
A:Introns: 23/3  
A:Keywords: blocked carboxyl end; glycoprotein; lipoprotein; membrane protein; phosphatidyl  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:27-56/Product: heat-stable antigen ML/69-JlId #status predicted <MAT>  
F:57-76/Domain: carboxyl-terminal propeptide #status predicted <CRP>  
F:57/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Ser) (in mature form)

Query Match 0.7%; Score 8; DB 2; Length 76;  
Best Local Similarity 100.0%; Pred. No. 6.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159  
|||||  
Db 12 GLLLLALL 19

RESULT 15  
I53107  
CD24 precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999  
C:Accession: I53107; S25146  
R.Shirasawa, T.; Akashi, T.; Sakamoto, K.; Takahashi, H.; Maruyama, N.; Hirokawa, K.  
Dev. Dyn. 198, 1-13, 1993  
A:Title: Gene expression of CD24 core peptide molecule in developing brain and developin  
A:Reference number: I53107; MUID:94122434; PMID:8292828  
A:Accession: I53107  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-76 <RES>  
A:Cross-references: EMBL:Z11663; NID:G55903; PIDN:CAA77731.1; PID:G55902  
C:Keywords: phosphatidylinositol linkage

Query Match 0.7%; Score 8; DB 2; Length 76;  
Best Local Similarity 100.0%; Pred. No. 6.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159  
|||||  
Db 12 GLLLLALL 19

RESULT 16  
A48996  
B cell surface antigen CD24 precursor - human  
N:Alternate names: cluster-w4 antigen; signal transducer CD24  
C:Species: Homo sapiens (man)  
C:Date: 13-Dec-1993 #sequence\_revision 02-Aug-1996 #text\_change 21-Jul-2000  
C:Accession: I56114; A48996; I54201; I72676  
R.Kay, R.; Rosten, P.M.; Humphries, R.K.  
J. Immunol. 147, 1412-1415, 1991  
A:Title: CD24, a signal transducer modulating B cell activation responses, is a very shd

A:Reference number: I56114; MUID:91332458; PMID:1831224  
A:Accession: I56114  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-80 <RES>  
A:Cross-references: GB:I33930; NID:G500848; PID:G500849; GB:M58664; NID:G180167; PID:  
R.Jackson, D.; Walbel, R.; Weber, E.; Bell, J.; Stahel, R.A.  
Cancer Res. 52, 5264-5270, 1992  
A:Title: CD24, a signal-transducing molecule expressed on human B cells, is a major s  
A:Reference number: A48996; MUID:93007871; PMID:1327504  
A:Accession: A48996  
A:Molecule type: mRNA  
A:Residues: 1-56, 'V', 58-80 <JAC>  
A:Cross-references: GB:X69397; GB:S44888; NID:G996167; PIDN:CAA49195.1; PID:G996168  
A:Experimental source: small cell lung carcinoma line SW2  
A>Note: sequence extracted from NCBI backbone (NCBIN:114635, NCBI:P:114636)  
A>Note: both 57-Val and 57-Ala were found in small cell carcinoma line DC571  
R.Hough, M.R.; Rosten, P.M.; Sexton, T.L.; Kay, R.; Humphries, R.K.  
Genomics 22, 154-161, 1994  
A:Title: Mapping of CD24 and homologous sequences to multiple chromosomal loci.  
A:Reference number: I54201; MUID:95048364; PMID:7959762  
A:Accession: I54201  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 5-11, 'W', 13-43, 'T', 45-80 <RE2>  
A:Cross-references: GB:S75311; NID:G833885; PIDN:AA14170.1; PID:G4261870  
C:Genetics:  
A:Gene: GDB:CD24  
A:Cross-references: GDB:383795; OMIM:600074  
A:Map position: 6q21-6q21  
C:Keywords: B-cell; blocked carboxyl end; glycoprotein; lipoprotein; phosphatidylinos  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-67/Product: B cell surface antigen CD24 #status predicted <MAT>  
F:68-80/Domain: carboxyl-terminal propeptide #status predicted <PRO>  
F:67/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Ser) (in mature fo

Query Match 0.7%; Score 8; DB 2; Length 80;  
Best Local Similarity 100.0%; Pred. No. 7.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159  
|||||  
Db 12 GLLLLALL 19

RESULT 17  
AF0856  
conserved hypothetical protein STY3056 [imported] - Salmonella enterica subsp. enteri  
C:Species: Salmonella enterica subsp. enterica serovar Typhi  
A>Note: this species has also been called Salmonella typhi  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C:Accession: AF0856  
R.Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church  
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AF0856  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-103 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD06037.1; PID:G16504004; GSPDB:GN00176  
C:Genetics:  
A:Gene: STY3056  
C:Superfamily: hypothetical protein HI0673

Query Match 0.7%; Score 8; DB 2; Length 103;  
Best Local Similarity 100.0%; Pred. No. 9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLALL 1160



```

Db      6 LLLALLLV 13
|||||
RESULT 18
B75301
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C:Accession: B75301
R:White, O.; Eisen, J.A.; Heidelberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: B75301
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-127 <WHI>
A:C:Cross-references: GB:AE002054; GB:AE000513; NID:g6460010; PIDN:AAF11763.1; PID:g646001
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2210
A:Map position: 1
C:Superfamily: Deinococcus radiodurans hypothetical protein DR2210
Query Match      0.7%; Score 8; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 11;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1152 GLLLLALL 1159
|||||
Db      7 GLLLLALL 14
|||||
RESULT 19
S15785
heat-stable antigen-related hypothetical protein HSA-C - mouse
C:Species: Mus musculus (house mouse)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C:Accession: S15785
R:Wenger, R.H.; Ayane, M.; Bose, R.; Koehler, G.; Nielsen, P.J.
Eur. J. Immunol. 21, 1039-1046, 1991
A:Title: The genes for a mouse hematopoietic differentiation marker called the heat-stab
A:Reference number: S15783; MUID:91209380; PMID:2019286
A:Accession: S15785
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-141 <WEN>
A:C:Cross-references: EMBL:X56486; NID:q51441; PIDN:CAA19843.1; PID:g51442
A:Note: the authors translated the codon TTC for residue 87 as Ser and AGA for residue 1
A:Note: the authors did not translate the codon for residue 101
C:Genetics:
A:Introns: #status absent
C:Keywords: glycoprotein; membrane protein; phosphatidylinositol linkage
Query Match      0.7%; Score 8; DB 2; Length 141;
Best Local Similarity 100.0%; Pred. No. 12;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1152 GLLLLALL 1159
|||||
Db      12 GLLLLALL 19
|||||
RESULT 20
E71228
hypothetical protein PH0087 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: E71228
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

```

```

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: E71228
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-224 <KAW>
A:C:Cross-references: GB:AP000001; NID:g3236128; PIDN:BA239156.1; PID:g3256473
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBa
C:Genetics:
A:Gene: PH0087
Query Match      0.7%; Score 8; DB 2; Length 224;
Best Local Similarity 100.0%; Pred. No. 18;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      512 RGKVVVYE 519
|||||
Db      105 RGKVVVYE 112
|||||
RESULT 21
S76800
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S76800
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima,
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76800
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-226 <KAN>
A:C:Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BAAL8712.1; PID:g165
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: conserved hypothetical protein MTH747
Query Match      0.7%; Score 8; DB 2; Length 226;
Best Local Similarity 100.0%; Pred. No. 18;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1073 RLVPNQEI 1080
|||||
Db      183 RLVPNQEI 190
|||||
RESULT 22
S15102
eosinophil major basic protein 2 precursor - guinea pig
C:Species: Cavia porcellus (guinea pig)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jul-2000
C:Accession: S15102; S18501
R:Aoki, I.; Shindeh, Y.; Nishida, T.; Nakai, S.; Hong, Y.M.; Mio, M.; Saito, T.; Tasa
FEBS Lett. 282, 56-60, 1991
A:Title: Comparison of the amino acid and nucleotide sequences between human and two
A:Reference number: S15102; MUID:91224343; PMID:2026266
A:Accession: S15102
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-234 <AOX>
A:C:Cross-references: DDBJ:D00817; NID:g3135095; PIDN:BA000697.1; PID:g2202933
A:Accession: S18501
A:Molecule type: protein
A:Residues: 116,'X',118-134,'X',136-137,'X',139-145;161-176;181-200 <AOX2>
C:Superfamily: eosinophil major basic protein precursor; C-type lectin homology
C:Keywords: antibiotic; chondroitin sulfate proteoglycan; cytotoxin; eosinophil; glyco

```

F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-115/Domain: propeptide #status predicted <PRO>  
F:116-234/Product: eosinophil major basic protein 2 #status experimental <MAT>  
F:117-232/Domain: C-type lectin homology <LCH>  
F:124,25/Binding site: carbohydrate (Ser) (covalent) #status predicted  
F:69/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted  
F:135-232,209-224/Disulfide bonds: #status predicted

Query Match 0.7%; Score 8; DB 1; Length 234;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 LLLALLV 1160  
|||||  
Db 4 LLLALLV 11

RESULT 23  
S22090  
catechol O-methyltransferase (EC 2.1.1.6) - rat  
N:Alternate names: catechol-O-methyltransferase  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 21-Jul-2000  
C:Accession: S22090; JQ0787; F52356  
R:Finonen, J.; Salminen, M.; Jalanko, A.; Ukkonen, S.; Ulmanen, I.  
A:Description: Structure of the rat catechol-O-methyltransferase gene. Separate promoter submitted to the EMBL Data Library, June 1992  
A:Reference number: S22090  
A:Accession: S22090  
A:Molecule type: DNA  
A:Residues: 1-264 <EN>  
A:Cross-references: EMBL:212651; NID:G55889; PIDN:CAA78276.1; PID:G55890  
R:Salminen, M.; Lundstroem, K.; Tilgmann, C.; Savolainen, R.; Kalkkinen, I.  
Gene 93, 241-247, 1990  
A:Title: Molecular cloning and characterization of rat liver catechol-O-methyltransferase  
A:Reference number: JQ0787; MUID:91033034; PMID:2227437  
A:Accession: JQ0787  
A:Molecule type: mRNA  
A:Residues: 44-264 <SAL>  
A:Cross-references: GB:M60753; NID:G203336; PIDN:AAA40881.1; PID:G203337  
A:Experimental source: liver  
R:Tenhunen, J.; Ulmanen, I.  
Biochem. J. 296, 595-600, 1993  
A:Title: Production of rat soluble and membrane-bound catechol O-methyltransferase forms  
A:Reference number: I52356; MUID:94107221; PMID:8280056  
A:Accession: I52356  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-264 <RES>  
A:Cross-references: EMBL:212651; NID:G55889; PIDN:CAA78276.1; PID:G55890  
C:Comment: This enzyme catalyzes the transfer of a methyl group from S-adenosyl-L-methionine to catechol  
C:Genetics: COMT  
A:Gene: COMT  
C:Superfamily: caffeoyl-CoA 3-O-methyltransferase  
C:Keywords: methyltransferase; S-adenosylmethionine; transmembrane protein

Query Match 0.7%; Score 8; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLALL 1159  
|||||  
Db 9 GLLALL 16

RESULT 24  
AG2637  
conserved hypothetical protein Atu0498 [imported] - Agrobacterium tumefaciens (strain C58)  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 19-Nov-2002  
C:Accession: AG2637  
R:Wood, D.W.; Scubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.  
C:Genetics: COMT  
A:Gene: COMT  
C:Superfamily: caffeoyl-CoA 3-O-methyltransferase  
C:Keywords: methyltransferase; S-adenosylmethionine; transmembrane protein

Query Match 0.7%; Score 8; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLALL 1159  
|||||  
Db 9 GLLALL 16

RESULT 24  
AG2637  
conserved hypothetical protein Atu0498 [imported] - Agrobacterium tumefaciens (strain C58)  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 19-Nov-2002  
C:Accession: AG2637  
R:Wood, D.W.; Scubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.  
C:Genetics: COMT  
A:Gene: COMT  
C:Superfamily: caffeoyl-CoA 3-O-methyltransferase  
C:Keywords: methyltransferase; S-adenosylmethionine; transmembrane protein

Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AG2637  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-308 <KUR>  
A:Cross-references: GB:AE008688; PIDN:AA41517.1; PID:G17738846; GSPDB:GN00186  
A:Experimental source: strain C58 (Dupont)  
C:Genetics: COMT  
A:Gene: Atu0498  
A:Map position: circular chromosome

Query Match 0.7%; Score 8; DB 2; Length 308;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1155 LLALLVLA 1162  
|||||  
Db 275 LLALLVLA 282

RESULT 25  
F97419  
BH1459 conserved hypothetical protein (AP001512) [imported] - Agrobacterium tumefaciens  
C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
C:Accession: F97419  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldm  
A:Authors: Liu, F.; Wollam, C.; Allinger, M.; Boughty, D.; Scott, C.; Lappas, C.; Markelz,  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium  
A:Reference number: A97359; MUID:21608551; PMID:11743194  
A:Accession: F97419  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-356 <KUR>  
A:Cross-references: GB:AB007869; PIDN:AAK86311.1; PID:G15155427; GSPDB:GN00169  
C:Genetics: COMT  
A:Gene: AGR\_C 881  
A:Map position: circular chromosome

Query Match 0.7%; Score 8; DB 2; Length 356;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1155 LLALLVLA 1162  
|||||  
Db 323 LLALLVLA 330

RESULT 26  
AG5016  
hypothetical protein AT4g01220 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: AG5016  
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp  
Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: AG5001; MUID:20083488; PMID:10617198  
A:Accession: AG5016  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-360 <STO>  
A:Cross-references: GB:NC\_001268; NID:G7267619; PIDN:CAB80931.1; GSPDB:GN00140  
C:Genetics: COMT  
A:Gene: AT4g01220  
A:Map position: 4

Query Match 0.7%; Score 8; DB 2; Length 360;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160  
|||||  
Db 44 LLLALLLV 51

RESULT 27  
S19172  
cytochrome P450 2B4 - rat (fragments)  
N;Alternate names: cytochrome P450 LM2  
N;Contains: oxidoreductase (EC 1.-.-.-)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 22-Nov-1993 #sequence\_revision 21-Jul-1995 #text\_change 04-Mar-2000  
C;Accession: S19172  
R;Yuan, P.M.; Ryan, D.E.; Levin, W.; Shively, J.E.  
Proc. Natl. Acad. Sci. U.S.A. 80, 1169-1173, 1983  
A;Title: Identification and localization of amino acid substitutions between two phenob  
A;Reference number: S19172; PMID:83144040; PMID:6572377  
A;Accession: S19172  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-158;159-200;201-310;311-367 <YUA>  
A;Experimental source: strain Long-Evans  
C;Genetics:  
A;Gene: CYP2B4  
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
C;Keywords: cytochrome; heme; iron; metalloprotein; microsome; monooxygenase; oxidore  
F;312/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 0.7%; Score 8; DB 2; Length 367;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160  
|||||  
Db 6 LLLALLLV 13

RESULT 28  
DH004  
dopamine receptor D4 - human  
C;Species: Homo sapiens (man)  
C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 02-Sep-1997  
C;Accession: S15079  
R;van Tol, H.H.M.; Bunzow, J.R.; Guan, H.C.; Sunahara, R.K.; Seeman, P.; Niznik, H.B.; C  
Nature 350, 610-614, 1991  
A;Title: Cloning of the gene for a human dopamine D(4) receptor with high affinity for d  
A;Reference number: S15079; PMID:91204054; PMID:1840645  
A;Accession: S15079  
A;Molecule type: DNA  
A;Residues: 1-387 <VAN>  
A;Cross-references: EMBL:X58497  
C;Genetics:  
A;Gene: DRD4  
A;Cross-references: GDB:127782; OMIM:126452  
A;Map position: 11p15.5-11p15.5  
A;Introns: 95/3; 133/2; 269/2; 321/1  
C;Superfamily: vertebrate rhodopsin  
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; neurotransmi  
F;34-60/Domain: transmembrane #status predicted <TM1>  
F;72-96/Domain: transmembrane #status predicted <TM2>  
F;110-131/Domain: transmembrane #status predicted <TM3>  
F;153-174/Domain: transmembrane #status predicted <TM4>  
F;192-214/Domain: transmembrane #status predicted <TM5>  
F;215-314/Domain: intracellular #status predicted <INT>  
F;315-339/Domain: transmembrane #status predicted <TM6>  
F;349-368/Domain: transmembrane #status predicted <TM7>  
F;3/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;108-185/Disulfide bonds: #status predicted  
F;149,239/Binding site: phosphate (Ser) (covalent) #status predicted

F;297,306/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 0.7%; Score 8; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLLV 1161  
|||||  
Db 81 LLLALLLV 88

RESULT 29  
S23860  
chloramphenicol resistance protein homolog opdE - Pseudomonas aeruginosa  
C;Species: Pseudomonas aeruginosa  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 31-Dec-2000  
C;Accession: S23860; H83368  
R;Huang, H.; Siehnell, R.J.; Bellido, P.; Rawling, E.; Hancock, R.E.W.  
submitted to the EMBL Data Library, July 1992  
A;Description: Analysis of two gene regions involved in the expression of the imipene  
A;Reference number: S23859  
A;Accession: S23860  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-402 <HUR>  
A;Cross-references: EMBL:Z14064; NID:945366; PIDN:CAA78446.1; PID:945368  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; I  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
A;Reference number: A82950; PMID:20437337; PMID:10984043  
A;Accession: H83368  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-402 <STO>  
A;Cross-references: GB:AE004648; GB:AE004091; NID:9948237; PIDN:AAG05607.1; GSPDB:GN  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: opdE; PA2219  
C;Superfamily: Streptomyces lividans chloramphenicol resistance protein  
C;Keywords: antibiotic resistance; transmembrane protein

Query Match 0.7%; Score 8; DB 1; Length 402;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1147 GSTIGELL 1154  
|||||  
Db 360 GSTIGELL 367

RESULT 30  
G83568  
probable permease of ABC transporter PA0605 [imported] - Pseudomonas aeruginosa (stra  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C;Accession: G83568  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; I  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
A;Reference number: A82950; PMID:20437337; PMID:10984043  
A;Accession: G83568  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-415 <STO>  
A;Cross-references: GB:AE004497; GB:AE004091; NID:9946478; PIDN:AAG03994.1; GSPDB:GN  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: PA0605

Query Match 0.7%; Score 8; DB 2; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLLA 1157  
 |||||  
 Db 388 LGGLLLLA 395

RESULT 31  
 AG3504  
 C:Species: Brucella melitensis (strain 16M)  
 C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 19-Apr-2002  
 C:Accession: AG3504  
 R:DelVecchio, V.G.; Kaparaj, V.; Redkar, R.J.; Patra, G.; Majer, C.; Los, T.; Ivanova,  
 ; Mazur, M.; Goltzman, E.; Selkov, E.; Elizek, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens  
 A:Reference number: -AD3252; PMID:11756688  
 A:Accession: AG3504  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-442 <KUR>  
 A:Cross-references: GB:AE008917; PIDN:AAL53202.1; PID:gl7984076; GSPDB:GN00190  
 A:Experimental source: strain 16M  
 C:Genetics:  
 A:Gene: BMEI2021  
 A:Map position: 1  
 C:Superfamily: folypolyglutamate synthase  
 C:Keywords: ligase

Query Match 0.7%; Score 8; DB 2; Length 442;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 615 NEDGLIDL 622  
 |||||  
 Db 241 NEDGLIDL 248

RESULT 32  
 T01721  
 C:Species: Arabidopsis thaliana (mouse-ear cross)  
 C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 24-Mar-1999  
 C:Accession: T01721  
 R:Scheet, P.; Maggi, L.  
 Submitted to the EMBL Data Library, June 1997  
 A:Description: The sequence of A. thaliana IG002N01.  
 A:Reference number: Z14407  
 A:Accession: T01721  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-444 <SCH>  
 A:Cross-references: EMBL:AF007269; NID:g2191126; PID:g2191131  
 A:Experimental source: cultivar Columbia  
 C:Genetics:  
 A:Map position: 4  
 A:Introns: 68/1; 235/3; 302/3; 374/3  
 A:Note: A\_IG002N01.8

Query Match 0.7%; Score 8; DB 2; Length 444;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160  
 |||||  
 Db 115 LLLALLLV 122

RESULT 33  
 T47107

benzaldehyde dehydrogenase (NAD) (EC 1.2.1.28) [validated] - Pseudomonas putida plasm  
 C:Species: Pseudomonas putida  
 C:Date: 17-Mar-2000 #sequence\_revision 17-Mar-2000 #text\_change 03-Jun-2002  
 C:Accession: I47107; S13388  
 R:Inoue, J.; Shaw, J.P.; Rexik, M.; Harayama, S.  
 J. Bacteriol. 177, 1196-1201, 1995  
 A:Title: Overlapping substrate specificity of benzaldehyde dehydrogenase (the xylC ge  
 Pseudomonas putida.  
 A:Reference number: Z24352; MUID:95173094; PMID:7868591  
 A:Accession: T47107  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-487 <INO>  
 A:Cross-references: EMBL:U15151; NID:g555994; PIDN:AAA66218.1; PID:g555995  
 R:Chalmers, R.M.; Keen, J.N.; Fawson, C.A.  
 Biochem. J. 273, 99-107, 1991  
 A:Title: Comparison of benzyl alcohol dehydrogenases and benzaldehyde dehydrogenases  
 ene pathway in Pseudomonas putida.  
 A:Reference number: S13398; MUID:9113163; PMID:1989592  
 A:Accession: S13398  
 A:Molecule type: protein  
 A:Residues: 1-44 <CHA>  
 A:Experimental source: strain MT53  
 C:Genetics:  
 A:Gene: xylC  
 A:Genome: plasmid  
 A:Note: TOL plasmid pMW0  
 C:Function:  
 A:Description: EC 1.2.1.28 [validated, MUID:95173094]  
 C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology  
 C:Keywords: oxidoreductase

Query Match 0.7%; Score 8; DB 2; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 626 ALGNVAVIL 633  
 |||||  
 Db 171 ALGNVAVIL 178

RESULT 34  
 O4RTPE  
 C:Species: cytochrome P450 b; cytochrome P450, phenobarbital-inducible  
 N:Alternate names: cytochrome P450 b; cytochrome P450, phenobarbital-inducible  
 C:Contains: unspecific monooxygenase (EC 1.14.14.1)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 18-Aug-1982 #sequence\_revision 17-May-1996 #text\_change 03-Mar-2000  
 C:Accession: A00176; A54251; A23363; A29298; S03854; A92255; I54796  
 R:Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 2793-2797, 1982  
 A:Title: Primary structure of a cytochrome P-450: coding nucleotide sequence of phenol  
 A:Reference number: A93912; MUID:8222224; PMID:6953431  
 A:Accession: A00176  
 A:Molecule type: mRNA  
 A:Residues: 6-491 <FUJ>  
 A:Cross-references: EMBL:J00719; NID:g203752; PIDN:AAA41024.1; PID:g203753  
 A:Note: the authors translated the codon GAT for residue 166 as Glu, CTC for residue  
 R:Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 5443, 1982  
 A:Title: Primary structure of a cytochrome P450: coding nucleotide sequence of phenob.  
 A:Reference number: A93925  
 A:Contents: annotation  
 A:Note: the mistranslations shown in reference A93912 are acknowledged  
 R:Roberts, B.S.; Hopkins, N.E.; Zaluzec, E.J.; Gage, D.A.; Alworth, W.L.; Hollenberg,  
 Biochemistry 31, 3766-3771, 1994  
 A:Title: Identification of active-site peptides from (3)H-labeled 2-ethynynaphthalen  
 A:Reference number: A54251; MUID:94190899; PMID:8142377  
 A:Accession: A54251  
 A:Molecule type: protein  
 A:Residues: 290-301, 'X' <ROB>  
 R:Suwa, Y.; Mizukami, Y.; Sogawa, K.; Fujii-Kuriyama, Y.  
 J. Biol. Chem. 260, 7980-7984, 1985

A;Title: Gene structure of a major form of phenobarbital-inducible cytochrome P-450 in z  
A;Reference number: A22363; MUID:95234490; PMID:2989270  
A;Accession: A22363  
A;Molecule type: DNA  
A;Residues: 1-91, 'P', 93-204, 'R', 206-327, 'V', 328-356, 'H', 358-391, 'R', 393-415, 'V', 417-433,  
A;Cross-references: GB:U00320; NID:G203816; PIDN:AAA41046.1; PID:G203818  
A;Note: the authors translated the codon CAG for residue 57 as Gly, CCG for residue 92 as  
as Glu, AAA for residue 236 as Leu, AGC for residue 259 as Asn, GTT for residue 328 as T  
as Arg  
R;Rangarajan, P.N.; Ravishanker, H.; Padmanaban, G.  
Biochem. Biophys. Res. Commun. 144, 258-263, 1987  
A;Title: Isolation of a cytochrome P-450 gene variant and characterization of its 5' fl  
A;Reference number: A29298; MUID:97213174; PMID:3579906  
A;Accession: A29298  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-57 <RAN>  
R;Oesch, F.; Waxman, D.J.; Morrissey, J.J.; Honscha, W.; Kissel, W.; Friedberg, T.  
Arch. Biochem. Biophys. 270, 23-32, 1989  
A;Title: Antibodies targeted against hypervariable and constant regions of cytochromes P  
A;Reference number: 803854; MUID:89192373; PMID:2539047  
A;Accession: 803854  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-18;146-160, 'E', 162-165;166,330-361;362-380;402-423 <OES>  
R;Botelho, L.H.; Ryan, D.E.; Levin, W.  
J. Biol. Chem. 254, 5635-5640, 1979  
A;Title: Amino acid compositions and partial amino acid sequences of three highly purified  
or 3-methylcholanthrene.  
A;Reference number: A92255; MUID:79194111; PMID:109438  
A;Accession: A92255  
A;Molecule type: protein  
A;Residues: 1-3, 'T', 5-22 <BOT>  
R;Fuji-Kuriyama, Y.; Mizukami, Y.; Taniguchi, T.; Muramatsu, M.  
Int. Symp. Princess Takamatsu Cancer Res. Fund 12, 31-40, 1982  
A;Title: Molecular cloning and coding nucleotide sequence of complementary DNA of cytoch  
A;Reference number: 154796; MUID:83160754; PMID:6300027  
A;Accession: 154796  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 6-491 <RES>  
A;Cross-references: GB:M37134; NID:G203784; PIDN:AAC42028.1; PID:G203785  
C;Genetics:  
A;Gene: CYP2B1  
A;Introns: 57/3; 112/1; 162/1; 215/3; 274/3; 322/1; 384/3; 432/1  
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
C;Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo  
P;295-458/Domain: cytochrome P450 homology <P45>  
P;302/Active site: Thr #status predicted  
P;436/Binding site: heme iron (Cys) (axial ligand) #status predicted  
Query Match 0.74; Score 8; DB 1; Length 491;  
Best Local Similarity 100.0%; Pred. No.35;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1153 LLLALLLV 1160  
Db 6 LLLALLLV 13  
RESULT 35  
O4RTP2  
Cytochrome P450 2B2 - rat  
N;Alternate names: cytochrome P450 PB-4; cytochrome P450, phenobarbital-inducible; cyto  
N;Contains: oxidoreductase (EC 1.-.-.-)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 04-Dec-1986 #sequence revision 17-May-1996 #text change 01-Dec-2000  
A;Accession: A21162; A00177; B00176; B92255; S15589; A21872; A32736; S03855; I59060  
R;Mizukami, Y.; Sogawa, K.; Suwa, Y.; Muramatsu, M.; Fujii-Kuriyama, Y.  
Proc. Natl. Acad. Sci. U.S.A. 80, 3958-3962, 1983  
A;Title: Gene structure of a phenobarbital-inducible cytochrome P-450 in rat liver.  
A;Reference number: A21162; MUID:83247397; PMID:6306654  
A;Accession: A21162

A;Molecule type: DNA  
A;Residues: 1-472, 'M', 474-491 <MIZ>  
A;Cross-references: EMBL:J00728; NID:G203845; PIDN:AAA1056.1; PID:G203847  
A;Note: the authors translated the codon AGT for residue 4 as Thr, and ATG for residue  
R;Frey, A.B.; Waxman, D.J.; Kreibich, G.  
J. Biol. Chem. 260, 15253-15265, 1985  
A;Title: The structure of phenobarbital-inducible rat liver cytochrome P-450 isoform  
A;Reference number: A00177; MUID:86059379; PMID:3877725  
A;Accession: A00177  
A;Molecule type: protein  
A;Residues: 1-291, 'P', 293-320, 'AE', 323-475, 'D', 477-491 <FRE>  
R;Fuji-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.  
Proc. Natl. Acad. Sci. U.S.A. 79, 2793-2797, 1982  
A;Title: Primary structure of a cytochrome P-450: coding nucleotide sequence of phen  
A;Reference number: A93912; MUID:82222224; PMID:6953431  
A;Accession: B00176  
A;Molecule type: mRNA  
A;Residues: 6-359, 'S', 361-362, 'V', 364-366, 'V', 368-406, 'S', 408-416, 'N', 418, 'A', 420-477  
A;Note: nucleotide sequence for residues 1-5 is not given  
A;Note: the authors translated the codon CAT for residue 166 as Glu, CTG for residue  
R;Fuji-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.  
Proc. Natl. Acad. Sci. U.S.A. 79, 5443, 1982  
A;Title: Primary structure of a cytochrome P450: coding nucleotide sequence of phenot  
A;Reference number: A93925  
A;Contents: annotation; revisions  
A;Note: the mistranslations in reference A93912 are acknowledged  
R;Botelho, L.H.; Ryan, D.E.; Levin, W.  
J. Biol. Chem. 254, 5635-5640, 1979  
A;Title: Amino acid compositions and partial amino acid sequences of three highly pur  
or 3-methylcholanthrene.  
A;Reference number: A92255; MUID:79194111; PMID:109438  
A;Accession: B92255  
A;Molecule type: protein  
A;Residues: 1-3, 'T', 5-22 <BOT>  
R;Lacroix, D.; Desrochers, M.; Lambert, M.; Anderson, A.  
Gene 86, 201-207, 1990  
A;Title: Alternative splicing of mRNA encoding rat liver cytochrome P450e (P450IIB2).  
A;Reference number: S15589; MUID:90215299; PMID:2323573  
A;Accession: S15589  
A;Molecule type: mRNA  
A;Residues: 105-113, 'F', 115-274, 'VSPAMRE', 275-321, 'E', 323-491 <LAC>  
A;Cross-references: EMBL:M34452; NID:G203679; PIDN:AAA41004.1; PID:G203680  
A;Note: translation of the nucleotide sequence is not complete  
R;Phillips, I.R.; Shephard, E.A.; Ashworth, A.; Rabin, B.R.  
Gene 24, 41-52, 1983  
A;Reference number: A21872  
A;Accession: A21872  
A;Molecule type: mRNA  
A;Residues: 168-321, 'E', 323-443, 'K', 445-491 <PHI>  
R;Affolter, M.; Anderson, A.  
Biochem. Biophys. Res. Commun. 118, 655-662, 1984  
A;Title: Segmental homologies in the coding and 3' non-coding sequences of rat liver  
A;Reference number: A32736; MUID:84153837; PMID:6322758  
A;Accession: A32736  
A;Molecule type: mRNA  
A;Residues: 385-491 <APP>  
A;Cross-references: GB:K01626; NID:G203782; PIDN:AAA41037.1; PID:G203783  
R;Oesch, F.; Waxman, D.J.; Morrissey, J.J.; Honscha, W.; Kissel, W.; Friedberg, T.  
Arch. Biochem. Biophys. 270, 23-32, 1989  
A;Title: Antibodies targeted against hypervariable and constant regions of cytochrome  
A;Reference number: S03854; MUID:89192373; PMID:2539047  
A;Accession: S03855  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 329-358, 'AS', 361;362,363-380;402-423 <OES>  
R;Atchison, M.L.; Adesnik, M.  
Proc. Natl. Acad. Sci. U.S.A. 83, 2300-2304, 1986  
A;Title: Gene conversion in a cytochrome P-450 gene family.  
A;Reference number: I59060; MUID:86205943; PMID:3458196  
A;Accession: I59060  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 323-431 <RES>

A:Cross-references: GB:M13234; NID:g203848; PIDN:AAA41057.1; PID:g554434  
 C:Genetics:  
 A:Introns: 384/3  
 C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
 C:Keywords: alternative splicing; chromoprotein; electron transfer; endoplasmic reticulum  
 F:295-458/Domain: cytochrome P450 homology <P45>  
 F:436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 0.7%; Score 8; DB 1; Length 491;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 LLLALLLV 1160  
 |||||  
 Db 6 LLLALLLV 13

## RESULT 36

B31047  
 testostosterone 16alpha-hydroxylase (EC 1.14.14.-) cytochrome P450 2B10 - mouse  
 N:Alternate names: cytochrome P450CDB; cytochrome P450pf3/46  
 C:Species: Mus musculus (house mouse)  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 04-Mar-2000

C:Accession: B31047; A60559  
 R:Noshiro, M.; Lakso, M.; Kawajiri, K.; Negishi, M.  
 Biochemistry 27, 6434-6443, 1988  
 A:Title: Rip locus: regulation of female-specific isozyme (I-P-450-16-alpha) of testoste  
 A:Reference number: A31047; MUID:89118235; PMID:3219345  
 A:Accession: B31047

A:Molecule type: mRNA  
 A:Residues: 1-500 <NO2>  
 A:Cross-references: EMBL:M21856; NID:g201968; PIDN:AAA04025.1; PID:g201969  
 A:Note: the authors translated the codon TTA for residue 55 as Phe, AGA for residue 133  
 as Val, AAA for residue 236 as Asn, AGG for residue 251 as Lys, AAG for residue 253 as  
 3 as Asp  
 A:Note: the authors translated the codon GCA for residue 281 as Thr, TTC for residue 283  
 8 as Arg, ACA for residue 321 as Ala, GTG for residue 331 as Leu, TCA for residue 349 as  
 S Ala, and GAC for residue 479 as Gly  
 A:Note: the sequence nucleotide translation from Fig. 8 is inconsistent with the nucleot  
 R:Bornheim, L.M.; Correia, M.A.  
 Mol. Pharmacol. 36, 377-383, 1989  
 A:Title: Purification and characterization of a mouse liver cytochrome P-450 induced by  
 A:Reference number: A60559; MUID:89384475; PMID:2779523

A:Accession: A60559  
 A:Molecule type: protein  
 A:Residues: 1-15 <BOR>  
 A:Note: this enzyme was induced by cannabidiol

C:Genetics:  
 A:Gene: Cyp2b-10  
 C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
 C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo  
 F:295-467/Domain: cytochrome P450 homology <P45>  
 F:445/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 0.7%; Score 8; DB 2; Length 500;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 LLLALLLV 1160  
 |||||  
 Db 6 LLLALLLV 13

## RESULT 37

AC1314  
 malolactic enzyme (malate dehydrogenase) homolog lmo1915 [imported] - Listeria monocytog  
 C:Species: Listeria monocytogenes  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001

A:Accession: AC1314  
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
 D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001  
 A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;  
 ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AC1314

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-547 <GLA>

A:Cross-references: GB:NC\_003210; PIDN:CAC99993.1; PID:g16411368; GSPDB:GN00177

A:Experimental source: strain 5GD-e

C:Genetics:

A:Gene: lmo1915

C:Superfamily: malate dehydrogenase (oxaloacetate-decarboxylating)

Query Match 0.7%; Score 8; DB 2; Length 547;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 266 MIVITDGE 273  
 |||||  
 Db 141 MIVITDGE 148

## RESULT 38

AC1686  
 malolactic enzyme (malate dehydrogenase) homolog lin2029 [imported] - Listeria innocu  
 C:Species: Listeria innocua  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001

A:Accession: AC1686

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;

ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AC1686

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-547 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC97359.1; PID:g16414530; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: lin2029

C:Superfamily: malate dehydrogenase (oxaloacetate-decarboxylating)

Query Match 0.7%; Score 8; DB 2; Length 547;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 266 MIVITDGE 273  
 |||||  
 Db 141 MIVITDGE 148

## RESULT 39

W14EB  
 E1 protein - bovine papillomavirus type 1  
 C:Species: bovine papillomavirus type 1  
 C:Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 28-Jul-2000

A:Accession: A03663

R:Chen, E.Y.; Howley, P.M.; Levinson, A.D.; Seeburg, P.H.

Nature 299, 529-534, 1982

A:Title: The primary structure and genetic organization of the bovine papillomavirus

A:Reference number: A93289; MUID:81012974; PMID:6289124

A:Accession: A03663

A:Molecule type: DNA

A:Residues: 1-605 <CHE>

A:Cross-references: GB:X02346; GB:J02044; GB:M24622; GB:X00473; NID:g60965; PIDN:CAB4

C:Superfamily: papillomavirus E1 protein

C:Keywords: early protein

Query Match 0.7%; Score 8; DB 1; Length 605;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 927 LAAGSDSN 934  
 |||||  
 Db 337 LAAGSDSN 344

RESULT 40  
 A72429  
 oligopeptide ABC transporter, periplasmic oligopeptide-binding protein - Thermotoga maritima  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
 C:Accession: A72429  
 R:Nelson, X.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.N.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.  
 Nature 399, 323-329, 1999  
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing  
 A:Reference number: A72200; PMID:10360571  
 A:Accession: A72429  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-606 <ARN>  
 A:Cross-references: GB:AE001690; GS:AE000512; NID:g4980496; PIDN:AA035125.1; PID:g4980511  
 A:Experimental source: strain MSB8  
 C:Genetics:  
 A:Gene: TM0031  
 C:Superfamily: dipeptide transport protein

Query Match 0.7%; Score 8; DB 2; Length 606;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161  
 |||||  
 Db 9 LLLALLVL 16

RESULT 41  
 A69845  
 Na+/H+ antiporter homolog yjbQ - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999  
 C:Accession: A69845  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertoni, C.; Bron, S.; Brouillette, S.; Bruschini, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chao, A.; Ehrlich, S.D.; Emerson, P.T.; Estian, K.D.; Brington, J.; Fabbet, C.; Ferrari, E.; Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, Tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; PMID:98044033; PMID:9384377  
 A:Accession: A69845  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-614 <KUN>  
 A:Cross-references: GB:Z99110; GB:AL009126; NID:g2633472; PIDN:CAB13021.1; PID:ell83184;  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yjbQ

Query Match 0.7%; Score 8; DB 2; Length 614;

Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVL 1160  
 |||||  
 Db 313 LLLALLVL 320

RESULT 42  
 W1WLB2  
 E1 protein - bovine papillomavirus type 2  
 C:Species: bovine papillomavirus type 2  
 C:Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 12-Jun-1998  
 C:Accession: C31169  
 R:Groff, D.E.; Mitra, R.; Lancaster, W.D.  
 Submitted to GenBank, May 1988  
 A:Reference number: A94519  
 A:Accession: C31169  
 A:Molecule type: DNA  
 A:Residues: 1-620 <GRO>  
 A:Cross-references: GB:M20219; GS:M19551; NID:g332996  
 C:Superfamily: papillomavirus E1 protein  
 C:Keywords: early protein; glycoprotein  
 F:72,109,173/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 0.7%; Score 8; DB 1; Length 620;  
 Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 927 LAAGSDSN 934  
 |||||  
 Db 352 LAAGSDSN 359

RESULT 43  
 AC0527  
 ferrichrome transport protein FhuB precursor STV0221 [imported] - Salmonella enterica  
 C:Species: Salmonella enterica subsp. enterica serovar Typhi  
 A:Note: this species has also been called Salmonella typhi  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C:Accession: AC0527  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Park, S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Park, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, A.; Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi  
 A:Reference number: AB0502; PMID:21534947; PMID:11677608  
 A:Accession: AC0527  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-685 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD01355.1; PID:g16501482; GSPDB:GN00176  
 C:Genetics:  
 A:Gene: STV0221  
 C:Superfamily: vitamin B12 transport protein btuC

Query Match 0.7%; Score 8; DB 2; Length 685;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVL 1160  
 |||||  
 Db 36 LLLALLVL 43

RESULT 44  
 T16306  
 hypothetical protein F40F4.7 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T16306  
 R:Wilson, R.

submitted to the EMBL Data Library, November 1995  
 A:Description: The sequence of *C. elegans* cosmid F40F4.  
 A:Reference number: Z18493

A:Accession: T16306  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-697 <WIL>  
 A:Cross-references: EMBL:U40420; NID:g1065513; PID:g1065515; PIDN:AAA81431.1; CESP:F40F4  
 C:Gene: CESP:F40F4.7  
 A:Introns: 14/2; 86/3; 135/3; 223/1; 244/1; 331/1; 365/1; 396/1; 435/1; 559/3; 627/2; 65

Query Match 0.7%; Score 8; DB 2; Length 697;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 557 YNDVVVGA 564  
 DB 593 YNDVVVGA 600

## RESULT 45

A88445  
 protein C26E6.10 [imported] - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
 A:Accession: A88445  
 R:Anonymous, The *C. elegans* Sequencing Consortium.  
 Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology  
 A:Reference number: A75000; MUID:99069613; PMID:9851916  
 A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: A88445  
 A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-739 <STO>  
 A:Cross-references: GB:chr\_III; PIDN:AAA21163.1; PID:g532810; GSPDB:GN00021; CESP:C26E6.  
 C:Gene: C26E6.10  
 A:Map position: 3

Query Match 0.7%; Score 8; DB 2; Length 739;

Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 430 TSVVSSRQ 437  
 DB 234 TSVVSSRQ 241

## RESULT 46

E83349  
 probable ClpA/B-type proteinase PA2371 [imported] - *Pseudomonas aeruginosa* (strain PA01)  
 C:Species: *Pseudomonas aeruginosa*  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: E83349

R:Stover, C.K.; Ham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
 ; Lory, S.; Olson, M.V.  
 Nature 405, 959-964, 2000

A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
 A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: E83349  
 A:Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-849 <STO>  
 A:Cross-references: GB:AE004663; GB:AE004091; NID:g9948405; PIDN:AG05759.1; GSPDB:GN001

A:Experimental source: strain PA01  
 C:Gene: PA2371  
 C:Superfamily: endopeptidase Clp ATP-binding chain

Query Match 0.7%; Score 8; DB 2; Length 849;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159  
 DB 111 GLLLLALL 118

## RESULT 47

T31437  
 integrin alpha chain SU2 - sea urchin (*Lytechinus variegatus*)  
 C:Species: *Lytechinus variegatus* (variegated urchin)  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jan-2000  
 C:Accession: T31437

R:Hertzler, P.L.; McClay, D.R.  
 submitted to the EMBL Data Library, May 1998  
 A:Description: Alpha SU2, a sea urchin integrin which binds laminin.

A:Reference number: Z21035  
 A:Accession: T31437  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
 A:Residues: 1-1041 <HER>  
 A:Cross-references: EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC3572.1

A:Experimental source: developmental stage embryo  
 C:Function: binds laminin  
 A:Description: binds laminin  
 C:Superfamily: integrin alpha-2b chain

Query Match 0.7%; Score 8; DB 2; Length 1041;  
 Best Local Similarity 100.0%; Pred. No. 68;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 498 LLVGAPMY 505  
 DB 308 LLVGAPMY 315

## RESULT 48

T25068  
 hypothetical protein T21C9.2 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T25068

R:McMurray, A.  
 submitted to the EMBL Data Library, May 1996  
 A:Reference number: Z19977

A:Accession: T25068  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA

A:Residues: 1-1065 <WIL>  
 A:Cross-references: EMBL:T273098; PIDN:CAA97331.1; GSPDB:GN00023; CESP:T21C9.2  
 A:Experimental source: clone T21C9

C:Gene: CESP:T21C9.2  
 A:Map position: 5  
 A:Introns: 67/2; 147/2; 207/3; 242/3; 293/3; 825/2; 897/1; 1032/2

Query Match 0.7%; Score 8; DB 2; Length 1065;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 866 FASLIQKE 873  
 DB 457 FASLIQKE 464

## RESULT 49

S40311  
 integrin - fruit fly (*Drosophila melanogaster*)  
 C:Species: *Drosophila melanogaster*  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 21-Jul-2000  
 C:Accession: S40311



R;Wehrli, M.; DiAntonio, A.; Fearnley, I.M.; Smith, R.J.; Wilcox, M.  
 Mech. Dev. 43, 21-36, 1993  
 A>Title: Cloning and characterization of alpha(P51), a novel Drosophila melanogaster integrin alpha chain  
 A;Reference number: S40311; MUID:94059764; PMID:8240969  
 A;Accession: S40311  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A;Residues: 1-1146 <WEH>  
 A;Cross-references: EMBL:X73975; NID:g440143; PIDN:CAA52155.1; PID:g440144  
 A;Gene: FlyBase:new  
 C;Genetics:  
 A;Cross-references: FlyBase:FBgn0004456  
 C;Superfamily: Integrin alpha-2b chain

Query Match 0.7%; Score 8; DB 2; Length 1146;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 8; Conservative 0; Mismatches 0; Gaps 0;  
 Indels 0;

Qy 1151 GGLLLALL 1158  
 |||||  
 Db 1094 GGLLLALL 1101

RESULT 50  
 RWHUIC  
 cell surface glycoprotein CD11b precursor [validated] - human  
 A;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Mac-1  
 A;Title: Integrin alpha chain; neutrophil adherence receptor alphaM chain  
 C;Species: Homo sapiens (man)  
 C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000  
 C;Accession: A31108; A26915; A41600; A30892; A32218; A46526; A26091; I52567  
 R;Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.  
 J. Biol. Chem. 263, 12403-12411, 1988  
 A;Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b)  
 A;Reference number: A28915; MUID:88257215; PMID:2454931  
 A;Accession: A28915  
 A:Molecule type: mRNA  
 A;Residues: 1-499,501-965, P. 967-1153 <ARN>  
 A;Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:G186935; PIDN:AAA594  
 A;Note: the authors translated the codon TAC for residue 1129 as Thr  
 A;Note: part of this sequence, including the amino end of the mature protein, was confirmed by R;Shelley, C.S.; Arnaout, M.A.  
 Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991  
 A;Title: The promoter of the CD11b gene directs myeloid-specific and developmentally regulated expression  
 A;Reference number: A41600; MUID:92073318; PMID:1683702  
 A;Accession: A41600  
 A:Molecule type: DNA  
 A;Residues: 1-9 <SHE>  
 A;Cross-references: GB:M76724; NID:G180018; PIDN:AAA58410.1; PID:G553215  
 R;Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988  
 A;Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesion molecule-1  
 A;Reference number: A94193; MUID:88190151; PMID:2833753  
 A;Accession: A30892  
 A:Molecule type: mRNA  
 A;Residues: 917-1042 <AR2>  
 A;Cross-references: GB:M18044  
 R;Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989  
 A;Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor  
 A;Reference number: A32218; MUID:89098893; PMID:2563162  
 A;Accession: A32218  
 A:Molecule type: mRNA

A;Residues: 9-1153 <HIC>  
 A;Cross-references: GB:J04145; NID:G189068; PIDN:AAA59903.1; PID:G386975  
 A;Note: part of this sequence was confirmed by protein sequencing  
 R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.  
 J. Immunol. 150, 480-490, 1993  
 A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alphaM chain during evolution  
 A;Reference number: A46526; MUID:93123748; PMID:8419480  
 A;Accession: A46526  
 A>Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A;Residues: 1-499,501-1153 <FLB>  
 A;Cross-references: GB:SS2227; NID:G263047; PIDN:AA524821.1; PID:G263049  
 A;Note: the last three bases of intron 13, CAG, are included in some but not all matings  
 R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.  
 Biochim. Biophys. Acta 874, 368-371, 1986  
 A;Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across species  
 A;Reference number: A90664; MUID:87076671; PMID:3539202  
 A;Accession: A26091  
 A:Molecule type: protein  
 A;Residues: 17-31 <PIE>  
 A;Experimental source: granulocytes  
 R;Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.  
 Blood 79, 865-870, 1992  
 A;Title: Characterization of the myeloid-specific CD11b promoter.  
 A;Reference number: I52567; MUID:92144986; PMID:1346576  
 A;Accession: I52567  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A;Residues: 1-9 <RES>  
 A;Cross-references: GB:M84477; NID:G180184; PIDN:AAA51960.1; PID:G553219  
 C;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1  
 C;Genetics:  
 A;Gene: GDB:ITGAM; CR3A  
 A;Cross-references: GDB:120599; OMIM:120980  
 A;Map position: 16p11.2-16p11.2  
 A;Note: promoter contains a GATA motif and two Sp1 consensus binding sites  
 C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat 1  
 C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer;  
 F17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>  
 F17-1108/Domain: extracellular #status predicted <EXT>  
 F148-318/Domain: von Willebrand factor type A repeat homology <VWA2>  
 F1465-473/Region: calcium/magnesium binding #status predicted  
 F530-538/Region: calcium/magnesium binding #status predicted  
 F593-601/Region: calcium/magnesium binding #status predicted  
 F1169-1134/Domain: transmembrane #status predicted <TM>  
 F1195-1153/Domain: intracellular #status predicted <INT>  
 P:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Bind

Query Match 0.7%; Score 8; DB 1; Length 1153;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1151 GGLLLALL 1158  
 |||||  
 Db 1116 GGLLLALL 1123

RESULT 51  
 RWHUIC  
 cell surface glycoprotein CD11c precursor - human  
 A;Alternate names: leukocyte adhesion receptor p150,95 alpha chain  
 C;Species: Homo sapiens (man)  
 C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 22-Jun-1999  
 C;Accession: A36584; A35543; S00864  
 R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.  
 J. Biol. Chem. 265, 12750-12751, 1990  
 A;Reference number: A36584  
 A;Contents: erratum  
 A;Accession: A36584  
 A:Molecule type: DNA

A:Residues: 1-1163 <COR>  
 A:Note: this revision to the sequence from reference A35543 includes the carboxyl end  
 R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.  
 J. Biol. Chem. 265, 2782-2786, 1990  
 A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.  
 A:Reference number: A35543; MUID:90153906; PMID:2303426  
 A:Accession: A35543  
 A:Molecule type: DNA  
 A:Residues: 1-834 <CO2>  
 A:Note: this sequence has been revised in reference A3584  
 R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.  
 EMBO J. 6, 4023-4028, 1987  
 A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte  
 A:Reference number: S00864; MUID:88166645; PMID:3327687  
 A:Accession: S00864  
 A:Molecule type: mRNA  
 A:Residues: 1-755, 'L', 757-1163 <CO3>  
 A:CROSS-references: GB:M81695; EMBL:Y00093; NID:9487829; PIDN:AAAS9180.1; PID:9487830  
 A:Note: part of this sequence was confirmed by protein sequencing  
 A:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my  
 C:Genetics:  
 A:Gene: GDB:ITGAX; CD11C  
 A:CROSS-references: GDB:119758; OMIM:151510  
 A:Map position: 16p11.2-16p11.2  
 C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homol  
 C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>  
 F:20-1107/Domain: extracellular #status predicted <EXT>  
 F:143-319/Domain: von Willebrand factor type A repeat homology <VWA4>  
 F:1108-1133/Domain: transmembrane #status predicted <TM>  
 F:1134-1163/Domain: intracellular #status predicted <INT>  
 F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match 0.7%; Score 8; DB 1; Length 1163;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1151 GGLLLAL 1158  
 |||||  
 Db 1115 GGLLLAL 1122

RESULT 52  
 A53213  
 Integrin alpha-B chain - human  
 C:Species: Homo sapiens (man)  
 C:Date: 19-Oct-1995 #sequence\_revision 31-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: A53213  
 R:Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.  
 J. Biol. Chem. 269, 6015-6025, 1994  
 A:Title: Molecular cloning of the human mucosal lymphocyte integrin alpha(B) subunit. Un  
 A:Reference number: A53213; MUID:94164962; PMID:8119947  
 A:Accession: A53213  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1179 <SHA>  
 A:CROSS-references: GB:L25851; NID:9457244; PID:9457245  
 C:Genetics:  
 A:Gene: GDB:ITGAE  
 A:CROSS-references: GDB:330801  
 A:Map position: 17p13  
 C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homol  
 F:199-371/Domain: von Willebrand factor type A repeat homology <VWA3>

Query Match 0.7%; Score 8; DB 2; Length 1179;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 380 VLLGAVGA 387  
 |||||  
 Db 416 VLLGAVGA 423

## RESULT 53

S57664  
 IGA-specific metalloendopeptidase homolog sepA precursor - Shigella flexneri  
 C:Species: Shigella flexneri  
 C:Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 08-Dec-2000  
 C:Accession: S57664; S69769; S69768  
 R:Benjelloun-Touimi, Z.; Sansonetti, P.J.; Parsot, C.  
 submitted to the EMBL Data Library, February 1995  
 A:Description: Characterization of SepA, the major extracellular protein of Shigella  
 A:Reference number: S57664  
 A:Accession: S57664  
 A:Molecule type: DNA  
 A:Residues: 1-1366 <BEN>  
 A:CROSS-references: EMBL:Z48219; NID:9886952; PIDN:CAA88252.1; PID:9886953  
 R:Benjelloun-Touimi, Z.; Sansonetti, P.J.; Parsot, C.  
 Mol. Microbiol. 17, 123-135, 1995  
 A:Title: SepA, the major extracellular protein of Shigella flexneri: autonomous secre  
 A:Reference number: S69768; MUID:96020667; PMID:7476198  
 A:Accession: S69769  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-66;1077-1366 <BEN>  
 A:CROSS-references: EMBL:Z48219; NID:9886952; PIDN:CAA88252.1; PID:9886953  
 A:Experimental source: strain M90T (serotype 5)  
 A:Accession: S69768  
 A:Molecule type: protein  
 A:Residues: 57-61, 'Q', 63-72, 'X', 544-557, 'X', 1057-1068 <BEP>  
 A:Note: 6-Glu, 14-Thr, 18-Gly and 26-Gln were also found  
 C:Genetics:  
 A:Gene: sepA  
 C:Superfamily: IGA-specific metalloendopeptidase  
 C:Keywords: extracellular protein  
 F:1-56/Domain: signal sequence #status predicted <SIG>  
 F:57-1366/Product: IGA-specific metalloendopeptidase homolog sepA #status predicted <

Query Match 0.7%; Score 8; DB 2; Length 1366;  
 Best Local Similarity 100.0%; Pred. No. 86;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 KDNRLGL 107  
 |||||  
 Db 1335 KDNRLGL 1342

## RESULT 54

D81675  
 Polymorphic membrane protein B/C family TC0695 [imported] - Chlamydia muridarum (stra  
 C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
 C:Accession: D81675  
 R:Read, T.B.; Bruhan, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.P.; White, O.; Hicke  
 C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39  
 A:Reference number: A81500; MUID:20150255; PMID:10684935  
 A:Accession: D81675  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1460 <TET>  
 A:CROSS-references: GB:AE002338; GB:AE002160; NID:97190724; PIDN:AAF39511.1; PID:9719  
 A:Experimental source: strain Nigg (MoPn)  
 C:Genetics:  
 A:Gene: TC0695

Query Match 0.7%; Score 8; DB 2; Length 1460;  
 Best Local Similarity 100.0%; Pred. No. 91;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 583 SILKTPKQ 590  
 |||||  
 Db 1311 SILKTPKQ 1318

RESULT 55  
 S67250  
 DNA-directed RNA polymerase (EC 2.7.7.6) I 190K chain - yeast (*Saccharomyces cerevisiae*)  
 N:Alternate names: DNA-directed RNA polymerase A 190K chain; protein 06276; protein YOR34  
 C:Species: *Saccharomyces cerevisiae*  
 C:Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 20-Jun-2000  
 C:Accession: S67250; A29926; S67396  
 R:Goffeau, A.; Purnelle, B.  
 submitted to the Protein Sequence Database, July 1996  
 A:Reference number: S67246  
 A:Accession: S67250  
 A:Molecule type: DNA  
 A:Residues: 1-1664 <GCF>  
 A:Cross-references: EMBL:Z75249; NID:G1420740; PIDN:CAA98665.1; PID:G1420741; MIPS:YOR34  
 R:Experimental source: strain S288C  
 R:Memet, S.; Gouy, M.; Marck, C.; Sentenac, A.; Buhler, J.M.  
 J. Biol. Chem. 263, 2830-2839, 1988  
 A:Title: RPA190, the gene coding for the largest subunit of yeast RNA polymerase A.  
 A:Reference number: A29926; MUID:3813933; PMID:2830265  
 A:Accession: A29926  
 A:Molecule type: DNA  
 A:Residues: 1-157, 159-1664 <MEM>  
 A:Cross-references: EMBL:J03530; NID:G172214; PIDN:AAA34890.1; PID:G172215  
 R:Purnelle, B.; Goffeau, A.  
 submitted to the EMBL Data Library, February 1996  
 A:Description: Nucleotide sequence analysis of a 40 kb segment on the right arm of yeast  
 chromosome I genes.  
 A:Reference number: S67392  
 A:Accession: S67396  
 A:Molecule type: DNA  
 A:Residues: 1-1664 <PUR>  
 A:Cross-references: EMBL:X95720; NID:G1199839; PIDN:CAA65029.1; PID:G1199844  
 C:Genetics:  
 A:Gene: SGD:RPA190  
 A:Cross-references: SGD:S0005868; MIPS:YOR341W  
 A:Map position: 15R  
 C:Superfamily: Trypanosoma DNA-directed RNA polymerase I largest chain  
 C:Keywords: DNA binding; nucleotidyltransferase; transcription; zinc finger  
 F:62-78/Region: zinc finger CCH motif  
  
 Query Match 0.7%; Score 8; DB 2; Length 1664;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 481 GSEITSVD 488  
 DB 3 GSEITSVD 15  
  
 RESULT 56  
 H88444  
 protein C26E6.12 [imported] - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
 C:Accession: H88444  
 R:anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology  
 A:Reference number: A75000; MUID:99069613; PMID:9851916  
 A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C\_ele  
 A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: H88444  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1802 <STO>  
 A:Cross-references: GB:chr\_III; PIDN:AAA21166.1; PID:G532813; GSPDB:GN00021; CESP:C26E6.  
 C:Genetics:  
 A:Gene: C26E6.12  
 A:Map position: 3  
  
 Query Match 0.7%; Score 8; DB 2; Length 1802;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 430 TSVVSSRQ 437  
 DB 1297 TSVVSSRQ 1304  
  
 RESULT 57  
 RRVUBY  
 genome polyprotein - bunyamwera virus  
 N:Alternate names: L protein  
 N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
 C:Species: bunyamwera virus  
 C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 11-Jun-1999  
 C:Accession: A33744  
 R:Elliot, R.M.  
 Virology 173, 426-436, 1989  
 A:Title: Nucleotide sequence analysis of the large (L) genomic RNA segment of Bunyamwera  
 A:Reference number: A33744; MUID:90085791; PMID:5596023  
 A:Accession: A33744  
 A:Molecule type: Genomic RNA  
 A:Residues: 1-2238 <ELL>  
 A:Cross-references: GB:X14383; NID:G58712; PIDN:CAA32553.1; PID:G58713  
 C:Genetics:  
 A:Map position: segment L  
 C:Superfamily: bunyamwera RNA-directed RNA polymerase  
 C:Keywords: nucleotidyltransferase  
  
 Query Match 0.7%; Score 8; DB 1; Length 2238;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 184 FLINILKK 191  
 DB 1333 FLINILKK 1340  
  
 RESULT 58  
 D81666  
 hypothetical protein TC0766 [imported] - *Chlamydia muridarum* (strain Nigg)  
 C:Species: *Chlamydia muridarum*, *Chlamydia trachomatis* MoPn  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
 C:Accession: D81666  
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hicke  
 C.; Dodson, R.; Gwin, W.; Nelson, W.; DeBoy, R.; McClarty, G.; Salzbe  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A:Title: Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39  
 A:Reference number: A81500; MUID:20150255; PMID:10684935  
 A:Accession: D81666  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-47 <TET>  
 A:Cross-references: GB:AE002345; GB:AE002160; NID:G7190791; PIDN:AAF39569.1; PID:G719  
 A:Experimental source: strain Nigg (MoPn)  
 C:Genetics:  
 A:Gene: TC0766  
  
 Query Match 0.6%; Score 7; DB 2; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1153 LLLALL 1159  
 DB 4 LLLALL 10  
  
 RESULT 59  
 T47376  
 hypothetical protein T5C2.40 - *Arabidopsis thaliana*  
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C:Accession: T47376  
 R:Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Lemcke, K.

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24463

A:Accession: T47376

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-74 <OBE>

A:Cross-references: EMBL:AL138664

A:Experimental source: cultivar Columbia; BAC clone T5C2

C:Genetics:

A:Map position: 3

A:Introns: 36/3; 55/3

A>Note: T5C2.40

Query Match 0.6%; Score 7; DB 2; Length 74;

Best Local Similarity 100.0%; Pred. No. 69;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 786 HCVPLV 792

|||||

Db 13 HCVPLV 19

RESULT 60

I49515

B144 protein B - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999

C:Accession: I49515

R:Tsuje, I.; Shen, P.

Immunogenetics 26, 378-380, 1987

A:Title: A gene in the H-2S: H-2D interval of the major histocompatibility complex which

A:Reference number: I49514; MUID:89031493; PMID:3117682

A:Accession: I49515

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-89 <RES>

A:Cross-references: GB:M18187; NID:g192097; PIDN:AAA37273.1; PID:g192099

Query Match 0.6%; Score 7; DB 2; Length 89;

Best Local Similarity 100.0%; Pred. No. 81;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1150 LGILLL 1156

|||||

Db 36 LGILLL 42

RESULT 61

A99774

probable excisionase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05099

C:Species: Escherichia coli

C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001

C:Accession: A99774

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: A99774

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-94 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA34584.1; PID:g13360621; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: EC81161

Query Match 0.6%; Score 7; DB 2; Length 94;

Best Local Similarity 100.0%; Pred. No. 85;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 637 PVQINA 643

|||||

Db 18 PVQINA 24

RESULT 62

F85636

probable excisionase for bacteriophage BP-933W xisW [imported] - Escherichia coli (st

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C:Accession: F85636

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: F85636

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-99 <STO>

A:Cross-references: GB:AF005174; NID:g12514277; PIDN:AAG55554.1; GSPDB:GN00145; UMGF:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: xisW

Query Match 0.6%; Score 7; DB 2; Length 99;

Best Local Similarity 100.0%; Pred. No. 89;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 637 PVQINA 643

|||||

Db 23 PVQINA 29

RESULT 63

G75257

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000

C:Accession: G75257

R:White, O.; Eissen, J.A.; Heideberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J

; M.; Shen, M.O.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: G75257

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-103 <WHI>

A:Cross-references: GB:AE002086; GB:AE000513; NID:g6460395; PIDN:AAF12117.1; PID:g6461

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR2573

A:Map position: 1

Query Match 0.6%; Score 7; DB 2; Length 103;

Best Local Similarity 100.0%; Pred. No. 92;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1157 ALLVIAL 1163

|||||

Db 15 ALLVIAL 21

RESULT 64

PH0141

integrin alpha chain - chicken (fragments)

C:Species: Gallus gallus (chicken)

C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 15-Sep-2003

C:Accession: PH0141

R:Syfrig, J.; Mann, K.; Paulsson, M.

Exp. Cell Res. 194, 165-173, 1991

A:Title: An abundant chick gizzard integrin is the avian alpha1-beta1 integrin heterox

A:Reference number: PH0141; MUID:91224161; PMID:1851093

A:Accession: P0141  
A:Molecule type: protein  
A:Residues: 1-104 <SYF>  
A:Experimental source: gizzard  
C:Comment: Integrins are a family of cellular adhesion receptors that mediate interaction  
C:Keywords: cell adhesion

Query Match 0.6%; Score 7; DB 2; Length 104;  
Best Local Similarity 100.0%; Pred. No. 93;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 QIGSYFG 481  
|||||  
Db 45 QIGSYFG 51

## RESULT 65

S57386

hypothetical protein orf 00954 - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae  
C:Date: 28-Oct-1995 #sequence\_revision 05-Sep-1996 #text\_change 20-Oct-2000  
C:Accession: S57386  
R:Zurawski, E.; Pearson, B.M.; Kalogeropoulos, A.; Schweizer, M.  
Yeast 11, 975-986, 1995  
A:Title: A 29.425 kb segment on the left arm of yeast chromosome XV contains more than 6  
A:Reference number: S57374; MUID:96021609; PMID:8533473  
A:Accession: S57386  
A:Status: nucleic acid sequence not shown; conceptual translation of pseudogene  
A:Molecule type: DNA  
A:Residues: 1-106 <ZUN>  
A:Cross-references: EMBL:X83121; NID:9600461; PIDN:CAA58196.1; PID:9600475  
C:Genetics: C:Comment: There is no evidence that this sequence is expressed.  
A:Map position: 15L  
C:Keywords: pseudogene

Query Match 0.6%; Score 7; DB 4; Length 106;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 SFLACSP 123  
|||||  
Db 17 SFLACSP 23

## RESULT 66

JC1127

major allergen chain 2 precursor, short form - cat  
C:Species: Felis silvestris catus (domestic cat)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 05-Nov-1999  
C:Accession: JC1127  
R:Griffith, I.J.; Craig, S.; Pollock, J.; Yu, X.B.; Morgenstern, J.P.; Rogers, B.L.  
Gene 113, 263-268, 1992  
A:Title: Expression and genomic structure of the genes encoding Fd1, the major allergen  
A:Reference number: JC1126; MUID:92241678; PMID:1572548  
A:Accession: JC1127  
A:Molecule type: DNA  
A:Residues: 1-107 <GRI>

A:Cross-references: GB:X62478; NID:9395406; PIDN:CAA44345.1; PID:9395407  
A:Experimental source: skin  
C:Genetics: A:Gene: Ch2  
A:Introns: 21/1; 81/3  
C:Keywords: glycoprotein  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:18-107/Product: major allergen chain 2, short form #status predicted <MAT>  
F:50/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 0.6%; Score 7; DB 2; Length 107;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1157 ALLVLAL 1163

Db 4 ALLVLAL 10  
|||||

## RESULT 67

CS6413

major allergen Fel d1 chain 2 precursor - cat  
C:Species: Felis silvestris catus (domestic cat)  
C:Date: 11-Aug-1995 #sequence\_revision 11-Aug-1995 #text\_change 01-Dec-2000  
C:Accession: CS6413; JC1145  
R:Morgenstern, J.P.; Griffith, I.J.; Brauer, A.M.; Rogers, B.L.; Bond, J.F.; Chapman,  
Proc. Natl. Acad. Sci. U.S.A. 88, 9690-9694, 1991  
A:Title: Amino acid sequence of Fel d1, the major allergen of the domestic cat: prote  
A:Reference number: A56413; MUID:92052157; PMID:1946388  
A:Accession: CS6413  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-109 <MOR>  
A:Cross-references: GB:M77341; NID:g163822; PIDN:AA041616.1; PID:g163823  
R:Griffith, I.J.; Craig, S.; Pollock, J.; Yu, X.B.; Morgenstern, J.P.; Rogers, B.L.  
Gene 113, 263-268, 1992  
A:Title: Expression and genomic structure of the genes encoding Fd1, the major allerg  
A:Reference number: JC1126; MUID:92241678; PMID:1572548  
A:Accession: JC1145  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 18-109 <GRI>  
A:Experimental source: salivary gland  
C:Keywords: glycoprotein  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:50/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 0.6%; Score 7; DB 2; Length 109;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1157 ALLVLAL 1163  
|||||

Db 4 ALLVLAL 10

## RESULT 68

INRB

insulin precursor - rabbit

N:Alternate names: preproinsulin

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 24-Apr-1984 #sequence\_revision 23-Aug-1997 #text\_change 18-Jun-1999

C:Accession: A53438; A01581

R:Devaskar, S.U.; Giddings, S.J.; Rajakumar, P.A.; Carnaghi, L.R.; Menon, R.K.; Zahm,

J. Biol. Chem. 269, 8445-8454, 1994

A:Title: Insulin gene expression and insulin synthesis in mammalian neuronal cells.

A:Reference number: A53438; MUID:94179230; PMID:8132571

A:Accession: A53438

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-110 <DEV>

A:Cross-references: GB:U03610; NID:g467970; PIDN:AAA19033.1; PID:g467971

R:Smith, L.F.

Am. J. Med. 40, 662-666, 1966

A:Title: Species variation in the amino acid sequence of insulin.

A:Reference number: A90029; MUID:66160119; PMID:5949593

A:Accession: A01581

A:Molecule type: protein

A:Residues: 25-54; 90-110 <SMI>

C:Superfamily: insulin

C:Keywords: hormone; pancreas

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-54/Domain: insulin chain B #status experimental <BCH>

F:57-87/Domain: connecting C peptide #status predicted <CPBP>

F:90-110/Domain: insulin chain A #status experimental <ACH>

F:31-96,43-109,95-100/Disulfide bonds: #status predicted

```
Query Match      0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1155 LLALLVL 1161
Db 10 LLALLVL 16

RESULT 69
IPRT1
insulin 1 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 23-Oct-1981 #sequence_revision 23-Oct-1991 #text_change 24-Sep-1999
C:Accession: A90788; A90789; A94231; B92120; I51945; A01589
C:CORDell, B.; Bell, G.; Tischer, E.; DeNoto, F.M.; Ullrich, A.; Rutter, W.J.
Cell 18, 533-543, 1979
A:Title: Isolation and characterization of a cloned rat insulin gene.
A:Reference number: A90788; XUID:80045034; PMID:498283
A:Accession: A90788
A:Molecule type: DNA
A:Residues: 1-110 <COR>
A:CROSS-references: GB:J00747; NID:G204956; PIDN:AAA41442.1; PID:G204957
R:Lomedico, P.; Rosenthal, N.; Efstratiadis, A.; Gilbert, W.; Kolodner, R.; Tizard, R.
Cell 18, 545-558, 1979
A:Title: The structure and evolution of the two nonallelic rat preproinsulin genes.
A:Reference number: A90789; XUID:80045035; PMID:498284
A:Accession: A90789
A:Molecule type: DNA
A:Residues: 1-110 <IOM>
A:CROSS-references: GB:J00747; NID:G204956; PIDN:AAA41442.1; PID:G204957
R:Steiner, D.F.; Clark, J.L.; Nolan, C.; Rubenstein, A.H.; Margoliash, E.; Aten, B.; Oye
Recent Prog. Horm. Res. 25, 207-282, 1969
A:Title: Proinsulin and the biosynthesis of insulin.
A:Reference number: A94231; XUID:70067613; PMID:4311938
A:Accession: A94231
A:Molecule type: protein
A:Residues: 25-54; 90-110 <STB>
R:Tager, H.S.; Steiner, D.F.
J. Biol. Chem. 247, 7936-7940, 1972
A:Title: Primary structures of the proinsulin connecting peptides of the rat and horse.
A:Reference number: A92120; XUID:73061498; PMID:4640931
A:Accession: B92120
A:Molecule type: protein
A:Residues: 57-87 <TAG>
R:Lomedico, P.T.; Rosenthal, N.; Kolodner, R.; Efstratiadis, A.; Gilbert, W.
Ann. N. Y. Acad. Sci. 343, 425-432, 1980
A:Title: The structure of rat preproinsulin genes.
A:Reference number: I51945; XUID:80240379; PMID:6249167
A:Accession: I51945
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-110 <RES>
A:CROSS-references: GB:M25584; NID:G204947; PIDN:AAA41439.1; PID:G204948
C:Genetics:
A:Gene: INS1
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <BCH>
F:25-54, 90-110/Product: insulin #status experimental <MAT>
F:57-87/Domain: connecting peptide #status experimental <CPEP>
F:90-110/Domain: insulin chain A #status experimental <ACH>
F:31-96, 43-109, 95-100/Disulfide bonds: #status experimental

Query Match      0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1155 LLALLVL 1161
Db 10 LLALLVL 16

RESULT 70
protein hdeA precursor - Escherichia coli (strain K-12)
N:Alternate names: 10K S protein precursor; hypothetical protein A
C:Species: Escherichia coli
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Mar-2002
C:Accession: S30268; S30266; S47730; A65149
R:Yoshida, T.; Ueguchi, C.; Yamada, H.; Mizuno, T.
Mol. Gen. Genet. 237, 113-122, 1993
A:Title: Function of the Escherichia coli nucleoid protein, H-NS: molecular analysis
A:Reference number: S30261; XUID:93204884; PMID:8455549
A:Accession: S30268
A:Molecule type: DNA
A:Residues: 1-110 <YOS>
A:CROSS-references: EMBL:D11109; NID:G216429; PIDN:BAAC1883.1; PID:G216431
A:Accession: S30266
A:Molecule type: protein
A:Residues: 22-40 <YOM>
R:Plunkett, G.
Submitted to the EMBL Data Library, March 1994
A:Reference number: S47666
A:Accession: S47730
A:Molecule type: DNA
A:Residues: 1-110 <PLU>
A:CROSS-references: EMBL:U00039; NID:G466582; PIDN:AAB18486.1; PID:G466647
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; XUID:97426617; PMID:9278503
A:Accession: A65149
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-110 <BLAT>
A:CROSS-references: GB:AE000427; GB:U00036; NID:G1789919; PIDN:AAC76535.1; PID:G17899
C:Genetics:
A:Gene: hdeA
C:Superfamily: Escherichia coli protein hdeA
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-110/Product: 10K-S protein #status experimental <MAT>

Query Match      0.6%; Score 7; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1150 LGGILL 1156
Db 9 LGGILL 15

RESULT 71
F91177
protein hdeA precursor - Escherichia coli (strain O157:H7, substrain RMD 0509952)
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 27-Nov-2001
C:Accession: F91177
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; XUID:21156231; PMID:11258796
A:Accession: F91177
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-110 <HAY>
A:CROSS-references: GB:BA000007; PIDN:BA937813.1; PID:G13363864; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECS4390
C:Superfamily: Escherichia coli protein hdeA
```

Query Match 0.6%; Score 7; DB 2; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 97;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLL 1156  
 |||||  
 DB 9 LGGLLLL 15

## RESULT 72

G86023  
 C:Species: Escherichia coli (strain O157:H7, substrain EDL933)  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 27-Nov-2001  
 C:Accession: G86023  
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: G86023  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-110 <STO>  
 A:Cross-references: GB:AE005174; NID:gl2518219; PIDN:AA058651.1; GSPDB:GN00145; UWGP:249  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: hdeA  
 C:Superfamily: Escherichia coli protein hdeA

Query Match 0.6%; Score 7; DB 2; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 97;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLL 1156  
 |||||  
 DB 9 LGGLLLL 15

## RESULT 73

C95870  
 C:Species: Sinorhizobium meliloti  
 C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
 C:Accession: C95870  
 R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo  
 A:Reference number: A95942; MUID:21396508; PMID:11481431  
 A:Accession: C95870  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-114 <XOR>  
 A:Cross-references: GB:AL591985; PIDN:CAC48627.1; PID:gl5140099; GSPDB:GN00167  
 A:Experimental source: strain 1021, megaplasmid pSymB  
 R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
 Peta, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.P.;  
 L.; Hyman, R.W.; Jones, T.  
 Science 283, 668-672, 2001  
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A:Reference number: A96039; MUID:21368234; PMID:11474104  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: Smb20236  
 A:Genome: plasmid

Query Match 0.6%; Score 7; DB 2; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 RLGLSLA 110

DB 88 RLGLSLA 94  
 |||||

## RESULT 74

D75434  
 C:Species: Deinococcus radiodurans (strain R1)  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
 C:Accession: D75434  
 R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: D75434  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-115 <WHI>  
 A:Cross-references: GB:AE001962; GB:AE000513; NID:96458855; PIDN:AAF10702.1; PID:9645  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR1125  
 A:Map position: 1

Query Match 0.6%; Score 7; DB 2; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALL 1159  
 |||||  
 DB 10 LLLALL 16

## RESULT 75

B81215  
 C:Species: Neisseria meningitidis  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Feb-2001  
 C:Accession: B81215; 381792  
 R;Rettlein, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,  
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.  
 Ric, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masiagnani, V.; Pizza, M.  
 Science 287, 1809-1815, 2000  
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;  
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A:Reference number: A81000; MUID:20175755; PMID:10710307  
 A:Accession: B81215  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-132 <TET>  
 A:Cross-references: GB:AE002386; GB:AE002098; NID:97225512; PIDN:AAF40742.1; PID:9722  
 A:Experimental source: serogroup B, strain MC58  
 R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo  
 ; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre  
 Nature 404, 502-506, 2000  
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491  
 A:Reference number: A81775; MUID:20222556; PMID:10761919  
 A:Accession: G81792  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-132 <PAR>  
 A:Cross-references: GB:AL162758; GB:AL157959; NID:97380672; PIDN:CAB85407.1; PID:9738  
 A:Experimental source: serogroup A, strain Z2491  
 C:Genetics:  
 A:Gene: NMB0291; NWA2196

Query Match 0.6%; Score 7; DB 2; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLA 1157

Db           |||||||  
          45 GGLLLA 51

Search completed: June 24, 2004, 18:05:04  
Job time : 32 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 17:57:31 ; Search time 18 Seconds  
(without alignments)  
3436.630 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 1188

Sequence: 1 MDLPRGLVAVALSMPGFT.....FRSARRRRPGDTPPKVLE 1188

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	911	76.7	1189	1	ITAH_HUMAN
2	15	1.3	285	1	ITAI_CHICK
3	15	1.3	1151	1	ITAI_HUMAN
4	15	1.3	1180	1	ITAI_RAT
5	13	1.1	1167	1	ITAG_HUMAN
6	11	0.9	1170	1	ITB2_BOVIN
7	11	0.9	1181	1	ITR2_HUMAN
8	9	0.8	312	1	CAH4_BOVIN
9	9	0.8	371	1	LEU3_SCHPO
10	9	0.8	1178	1	ITR2_MOUSE
11	8	0.7	76	1	CD24_MOUSE
12	8	0.7	76	1	CD24_RAT
13	8	0.7	80	1	CD24_HUMAN
14	8	0.7	97	1	LST1_HUMAN
15	8	0.7	103	1	FTSB_SALTY
16	8	0.7	105	1	FTSB_KLEAE
17	8	0.7	179	1	IL22_HUMAN
18	8	0.7	234	1	EMB2_CAVPO
19	8	0.7	264	1	COMT_RAT
20	8	0.7	402	1	ORDE_PSEAE
21	8	0.7	423	1	NSMA_HUMAN
22	8	0.7	444	1	ENGA_MYCPE
23	8	0.7	451	1	HISX_COREF
24	8	0.7	467	1	D4DR_HUMAN
25	8	0.7	467	1	SIL7_HUMAN
26	8	0.7	470	1	CPBK_MOUSE
27	8	0.7	487	1	XILC_PSEPU
28	8	0.7	491	1	CPB1_RAT
29	8	0.7	491	1	CPB2_RAT
30	8	0.7	500	1	CPBA_MOUSE
31	8	0.7	604	1	VE1_BPV2
32	8	0.7	605	1	VE1_BPV1
33	8	0.7	811	1	TMS6_HUMAN

34	8	0.7	1050	1	ITAS_XENLA
35	8	0.7	1146	1	ITAM_DROME
36	8	0.7	1152	1	ITAM_HUMAN
37	8	0.7	1157	1	DD37_HUMAN
38	8	0.7	1162	1	ITAD_HUMAN
39	8	0.7	1163	1	ITAX_HUMAN
40	8	0.7	1173	1	TSP1_XENLA
41	8	0.7	1179	1	ITAE_HUMAN
42	8	0.7	1460	1	MPPC_CHLMU
43	8	0.7	1507	1	SET2_CABEL
44	8	0.7	1664	1	RPAL_YEAST
45	8	0.7	2238	1	RRPL_BUNYV
46	7	0.6	95	1	LST1_MOUSE
47	7	0.6	109	1	FEL2_FELCA
48	7	0.6	110	1	HDEA_ECOLI
49	7	0.6	110	1	INS1_RAT
50	7	0.6	110	1	INS_RABIT
51	7	0.6	139	1	IGF_MXGL
52	7	0.6	139	1	LAMP_PETMA
53	7	0.6	141	1	HBAD_CHICK
54	7	0.6	141	1	HBAD_MELGA
55	7	0.6	142	1	KK_CHICK
56	7	0.6	160	1	BIK_HUMAN
57	7	0.6	160	1	NUOE_AQUAE
58	7	0.6	161	1	TCH2_ARATH
59	7	0.6	171	1	YF87_METJA
60	7	0.6	176	1	YD2F_SCHPO
61	7	0.6	194	1	HIS7_CHLTE
62	7	0.6	201	1	Z239_MOUSE
63	7	0.6	206	1	GPBB_MOUSE
64	7	0.6	208	1	PSMB_SULSO
65	7	0.6	213	1	KPTA_PYPAE
66	7	0.6	214	1	CH13_HUMAN
67	7	0.6	214	1	CYB_CERCE
68	7	0.6	220	1	FGF3_CHICK
69	7	0.6	221	1	SDFL_HUMAN
70	7	0.6	223	1	EMBP_MOUSE
71	7	0.6	227	1	EMBP_RAT
72	7	0.6	233	1	EMBA_CAVPO
73	7	0.6	235	1	YBEU_ECOLI
74	7	0.6	235	1	YOGA_ECOLI
75	7	0.6	236	1	RR2_MAIZE
76	7	0.6	236	1	RR2_ORYSA
77	7	0.6	236	1	RR2_WHRAT
78	7	0.6	240	1	IBP6_HUMAN
79	7	0.6	253	1	GX42_RAT
80	7	0.6	253	1	Y685_METJA
81	7	0.6	260	1	COX3_ASTPE
82	7	0.6	260	1	COX3_PISOC
83	7	0.6	261	1	COX3_DINSE
84	7	0.6	261	1	COX3_PARLI
85	7	0.6	262	1	COX3_BRALA
86	7	0.6	262	1	COX3_METSE
87	7	0.6	263	1	LPSA_PASHA
88	7	0.6	270	1	REP8_HUMAN
89	7	0.6	276	1	NO5Y_PSEST
90	7	0.6	276	1	UPK_EACUS
91	7	0.6	295	1	YZ34_METJA
92	7	0.6	301	1	PANE_PYRHO
93	7	0.6	306	1	COAA_STPMU
94	7	0.6	309	1	NAT1_HUMAN
95	7	0.6	309	1	NTCB_SYNP7
96	7	0.6	311	1	MRAW_CAUCR
97	7	0.6	311	1	PYDA_ENTFA
98	7	0.6	312	1	O8H2_HUMAN
99	7	0.6	313	1	LEU3_NEILA
100	7	0.6	315	1	CALU1_HUMAN

#### ALIGNMENTS

RESULT 1

ITAH\_HUMAN  
 ID ITAH\_HUMAN STANDARD; PRT; 1189 AA.  
 AC Q9UKX5; Q9UKQ1;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Integrin alpha-11 precursor.  
 GN ITGAL1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal heart, and Osteoblast;  
 RX MEDLINE=99417679; PubMed=10486209;  
 RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,  
 RA Wang S.-X., Morris C.M., Krissansen G.W.;  
 RT "Cloning, sequence analysis, and chromosomal localization of the novel  
 RT human integrin alpha11 subunit (ITGAL1).";  
 RL Genomics 60:179-187(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal muscle, and Uterus;  
 RX MEDLINE=99395147; PubMed=10464311;  
 RA Velling T., Kusche-Gullberg M., Seijersen T., Gullberg D.;  
 RT "cDNA Cloning and Chromosomal Localization of Human alpha(11)  
 RT Integrin. A collagen-binding, i domain-containing, beta(1)-associated  
 RT integrin alpha-chain present in muscle tissues.";  
 RL J. Biol. Chem. 274:25735-25742(1999).  
 RN [3]  
 RP SEQUENCE OF 954-1188 FROM N.A.  
 RC TISSUE=Fibroblast;  
 RA Andreu N., Estivill X., Escarceller M., Sumoy L.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.  
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11  
 CC ASSOCIATES WITH BETA-1.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND  
 CC HEART, INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO  
 CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO  
 CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS, IN  
 CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN  
 CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.  
 CC -!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING  
 CC FETAL MUSCLE CELLS (IN VITRO).  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; AF109681; AAF01258.1; --  
 DR EMBL; AF137378; AAD51919.2; --  
 DR EMBL; AL358064; CAB94392.1; --  
 DR HSSP; P17301; IAOX.  
 DR Genew; HGNC:6136; ITGAL1.  
 DR MIM; 604789; --  
 DR GO; GO:0008305; C:integrin complex; TAS.  
 DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.  
 DR GO; GO:0005518; F:collagen binding; TAS.  
 DR GO; GO:0007160; P:cell-matrix adhesion; TAS.  
 DR GO; GO:0007517; P:muscle development; TAS.  
 DR InterPro; IPR000413; Integrin\_alpha.

DR InterPro; IPR002035; VWFA\_A.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00092; vwa; 1.  
 DR PRINTS; PR01185; INTEGRINA.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00131; Int\_alpha; 5.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; FALSE\_NEG.  
 DR PROSITE; PS50234; VWFA; 1.  
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Signal; Repeat; Calcium; Magnesium; Polymorphism.  
 FT SIGNAL 1 22  
 FT CHAIN 23 1189  
 FT DOMAIN 23 1142  
 FT TRANSMEM 1143 1165  
 FT DOMAIN 1166 1189  
 FT REPEAT 38 94  
 FT REPEAT 102 163  
 FT DOMAIN 167 345  
 FT REPEAT 359 420  
 FT REPEAT 422 475  
 FT REPEAT 477 537  
 FT REPEAT 539 598  
 FT REPEAT 601 653  
 FT DOMAIN 1154 1162  
 FT DOMAIN 1174 1177  
 FT CA\_BIND 488 496  
 FT CA\_BIND 551 559  
 FT CA\_BIND 613 621  
 FT DISULFID 76 83  
 FT DISULFID 121 139  
 FT DISULFID 129 159  
 FT DISULFID 659 668  
 FT DISULFID 674 729  
 FT DISULFID 781 787  
 FT DISULFID 881 893  
 FT CARBOHYD 82 82  
 FT CARBOHYD 95 95  
 FT CARBOHYD 291 291  
 FT CARBOHYD 331 331  
 FT CARBOHYD 358 358  
 FT CARBOHYD 449 449  
 FT CARBOHYD 462 462  
 FT CARBOHYD 528 528  
 FT CARBOHYD 642 642  
 FT CARBOHYD 694 694  
 FT CARBOHYD 857 857  
 FT CARBOHYD 894 894  
 FT CARBOHYD 973 973  
 FT CARBOHYD 1032 1032  
 FT CARBOHYD 1040 1040  
 FT VARIANT 433 433  
 FT VARIANT 524 524  
 FT VARIANT 972 972  
 FT VARIANT 1003 1003  
 FT VARIANT 1030 1030  
 FT VARIANT 1094 1094  
 FT SEQUENCE 1189 AA; 133609 MW; 60303C08A4A4CD52 CRC64;  
 Query Match 76.7%; Score 911; DB 1; Length 1189;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDLPRGLVAVWALSFWPGFTDTFMNDRKPRVTPGSRTPAFGKTVQQHDSGNKWLVVGA 60  
 DB 1 MDLPRGLVAVWALSFWPGFTDTFMNDRKPRVTPGSRTPAFGKTVQQHDSGNKWLVVGA 60

QY 61 PLENGYQKTDVYKCPVHGNCTKLNIGRVTLNSVSRKDNVRLGLSLATNPKNISFLA 120  
 DB 61 PLENGYQKTDVYKCPVHGNCTKLNIGRVTLNSVSRKDNVRLGLSLATNPKNISFLA 120  
 QY 121 CSPLWSHSCGSSYYTTCGSRVNSFRSKTVAPALQCCOTYMDIVIVLDGNSIYPWVE 180  
 DB 121 CSPLWSHSCGSSYYTTCGSRVNSFRSKTVAPALQCCOTYMDIVIVLDGNSIYPWVE 180  
 QY 181 VOHPLINILKFFYIGFGQIOGVVQYGEDVVEHFLNDYRSVKDVEAASHIEQGGTET 240  
 DB 181 VOHPLINILKFFYIGFGQIOGVVQYGEDVVEHFLNDYRSVKDVEAASHIEQGGTET 240  
 QY 241 RTAGIEFARAEAFQKGRKAKKVMIVITGESHSDSPLEKVIQOSRDNVTYAVAVL 300  
 DB 241 RTAGIEFARAEAFQKGRKAKKVMIVITGESHSDSPLEKVIQOSRDNVTYAVAVL 300  
 QY 301 GYNNRGINPTFFNEIKYIASDDPKHFFNVVDEAALKDVIDALGDRIFSLGNTKNET 360  
 DB 301 GYNNRGINPTFFNEIKYIASDDPKHFFNVVDEAALKDVIDALGDRIFSLGNTKNET 360  
 QY 361 SFGLMSOTGFSHVVEDVLLGAVGYDMNGAVLKTSAKVIPLRESYIKKPEBELKN 420  
 DB 361 SFGLMSOTGFSHVVEDVLLGAVGYDMNGAVLKTSAKVIPLRESYIKKPEBELKN 420  
 QY 421 HGAVLGYTVTSVWSRQGRVYVAGAPRNHTGKVLFTMNNRSLTIHOAMRGQIGSYF 480  
 DB 421 HGAVLGYTVTSVWSRQGRVYVAGAPRNHTGKVLFTMNNRSLTIHOAMRGQIGSYF 480  
 QY 481 GSEITSVDIDGQVTVLLGAPMYFNEGRGKVVYVYELQNRFPVYNGTLKDSHSYQNA 540  
 DB 481 GSEITSVDIDGQVTVLLGAPMYFNEGRGKVVYVYELQNRFPVYNGTLKDSHSYQNA 540  
 QY 541 RFGSSIASVRDLNDSYNDVVGAPLEDNHAGALYIFHGPGRSILKTPKQITASELATG 600  
 DB 541 RFGSSIASVRDLNDSYNDVVGAPLEDNHAGALYIFHGPGRSILKTPKQITASELATG 600  
 QY 601 LQYFGCSIHGQDLNEDGLIDLAVGALNAVILSRPVVQINASLHFEPSKINIFHRDCK 660  
 DB 601 LQYFGCSIHGQDLNEDGLIDLAVGALNAVILSRPVVQINASLHFEPSKINIFHRDCK 660  
 QY 661 RSGRDATCLAFCTPIPLAPHQTTTGVRYNATMDERRYTPRAHLDEGGDFTNRVAV 720  
 DB 661 RSGRDATCLAFCTPIPLAPHQTTTGVRYNATMDERRYTPRAHLDEGGDFTNRVAV 720  
 QY 721 LLSGQELCERINFHVLDTADYVRPVSVEYSLEDPDGPMLDDGWPPTLLRVSVPPWNG 780  
 DB 721 LLSGQELCERINFHVLDTADYVRPVSVEYSLEDPDGPMLDDGWPPTLLRVSVPPWNG 780  
 QY 781 CNEDEHCVDPDLVLDARSOLPTAMEYQORVLRKPQDCSAYTSLSDTTVFIIESTRQVAV 840  
 DB 781 CNEDEHCVDPDLVLDARSOLPTAMEYQORVLRKPQDCSAYTSLSDTTVFIIESTRQVAV 840  
 QY 841 EATLENGENAYSTVLNISQSANLQFASLIQKESDGSIECVNEERLQKQVCNVSYPFF 900  
 DB 841 EATLENGENAYSTVLNISQSANLQFASLIQKESDGSIECVNEERLQKQVCNVSYPFF 900  
 QY 901 RAKAKVAFRLD 911  
 DB 901 RAKAKVAFRLD 911

RESULT 2

ITAL CHICK STANDARD; PRT; 285 AA.  
 AC Q90615;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a)  
 DE (Fragment).  
 GN ITGAL.  
 OS Gallus Gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Gizzard;  
 RX MEDLINE=94357930; PubMed=7521332;  
 RT Kern A., Brisevitz R., Bank I., Marcantonio E.E.;  
 "The role of the I domain in ligand binding of the human integrin  
 alpha 1 beta 1.";  
 J Biol. Chem. 269:22811-22816 (1994).  
 CC -I- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND  
 COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-  
 Z-R IN COLLAGEN.  
 CC -I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1  
 ASSOCIATES WITH BETA-1.  
 CC -I- SUBCELLULAR LOCATION: TYPE I membrane protein.  
 CC -I- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -I- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -I- SIMILARITY: Contains 1 VWFA domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; U10114; AAA59067.1; -;  
 DR HSP; U17301; IAOX.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWFA.  
 DR Pfam; PF00092; vwa; 1.  
 DR SMART; SM00327; VWFA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; PARTIAL.  
 DR PROSITE; PS00234; VWFA; 1.  
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane.  
 FT NON TER 1 1 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN <1 285 VWFA.  
 FT CARBOHYD 2 2 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT NON TER 285 285  
 SQ SEQUENCE 285 AA; 31503 MW; 1B05D3246CD5CATE CRC64;  
 Query Match 1.3%; Score 15; DB 1; Length 285;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 164 DIVIVLDGNSIYPW 178  
 DB 63 DIVIVLDGNSIYPW 77  
 RESULT 3  
 ITAL HUMAN STANDARD; PRT; 1151 AA.  
 ID -ITAL HUMAN  
 AC P56139;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).  
 GN ITGAL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=93155124; PubMed=8428973;  
 RA Sriesewitz R., Epstein M.R., Marcantonio E.S.;  
 RT "Expression of native and truncated forms of the human integrin alpha  
 J. Biol. Chem. 268:2989-2996(1993).  
 RL J. Biol. Chem. 268:2989-2996(1993).  
 CC -I- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND  
 COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-  
 E-R IN COLLAGEN.  
 CC -I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1  
 ASSOCIATES WITH BETA-1.  
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -I- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -I- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -I- SIMILARITY: Contains 1 VWFA domain.  
 CC -I- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -I- DATABASE: NAME=PROW; NOTE=CD guide CD49a entry;  
 WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd49a.htm".  
 DR PIR; A45226;  
 DR PDB; 1QCS; 17-MAY-00.  
 DR Genew; HGNC:6134; ITGAI.  
 DR MIM; 192968;  
 DR GO; GO:0008305; C: integrin complex; TAS.  
 DR GO; GO:0004895; F: cell adhesion receptor activity; NAS.  
 DR GO; GO:0005518; F: collagen binding; TAS.  
 DR GO; GO:0007160; P: cell-matrix adhesion; NAS.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWFA.  
 DR Pfam; PF01839; FG-GAP 3.  
 DR Pfam; PF00357; integrin\_A; 1.  
 DR Pfam; PF00092; vwa; 1.  
 DR SMART; SM00191; Int\_alpha; 5.  
 DR SMART; SM00327; VWFA; 1.  
 DR PROSITE; PS00234; VWFA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 Repeat; Calcium; Magnesium; 3D-structure.  
 FT DOMAIN 1 1113  
 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1114 1136  
 POTENTIAL.  
 FT DOMAIN 1137 1151  
 CYTOPLASMIC (POTENTIAL).  
 FT REPEAT 16 75  
 FG-GAP 1.  
 FT REPEAT 2 75  
 FG-GAP 2.  
 FT DOMAIN 147 360  
 VWFA.  
 FT REPEAT 349 404  
 FG-GAP 3.  
 FT REPEAT 405 457  
 FG-GAP 4.  
 FT REPEAT 459 520  
 FG-GAP 5.  
 FT REPEAT 540 599  
 FG-GAP 6.  
 FT REPEAT 602 654  
 FG-GAP 7.  
 FT CA\_BIND 470 478  
 POTENTIAL.  
 FT CA\_BIND 552 560  
 POTENTIAL.  
 FT CA\_BIND 614 622  
 POTENTIAL.  
 FT SITE 1139 1142  
 GPPAR MOTIF.  
 FT DISULFID 54 64  
 BY SIMILARITY.  
 FT DISULFID 660 669  
 BY SIMILARITY.  
 FT DISULFID 675 728  
 BY SIMILARITY.  
 FT DISULFID 780 786  
 BY SIMILARITY.  
 FT DISULFID 850 858  
 BY SIMILARITY.  
 FT DISULFID 1002 1034  
 BY SIMILARITY.  
 FT DISULFID 1037 1044  
 BY SIMILARITY.  
 FT CARBOHYD 46 46  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 72 72  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 77 77  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 84 84  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 189 189  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 289 289  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 313 313  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 374 374  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 390 390  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 432 432  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 504 504  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 671 671  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 720 720  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 752 752  
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 812 812  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 855 855  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 880 880  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 887 887  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 911 911  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 938 938  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 946 946  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 980 980  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1045 1045  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1055 1055  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1074 1074  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1085 1085  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1151 AA; 127837 MW; 683F3CIAABP52808 CRC64;  
 Query Match 1.3%; Score 15; DB 1; Length 1151;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-06; Indels 0; Gaps 0;  
 Matches 15; Conservative 0; Mismatches 0;  
 QY 164 DIVIVLDGNSIYPW 178  
 |||||  
 DB 144 DIVIVLDGNSIYPW 159  
 |||||  
 RESULT 4  
 ID ITAL RAT  
 AC P18614; STANDARD; PRT; 1180 AA.  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Integrin alpha-1 precursor (Laminin and collagen receptor) (VLA-1)  
 DE (CD49a).  
 GN ITGAI.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Theria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RX MEDLINE=90338125; PubMed=2380249;  
 RA Ignatius M.J., Large T.H., Houde M., Tawil J.W., Barton A.,  
 RA Esch P., Carbonetto S., Reichardt L.F.;  
 RT "Molecular cloning of the rat integrin alpha 1-subunit: a receptor  
 for laminin and collagen."  
 RL J. Cell Biol. 111:709-720(1990).  
 RN [2]  
 X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 151-364.  
 RX MEDLINE=99313197; PubMed=10386626;  
 RA Nolte M., Pepinsky R.B., Vanyaminov S.Y., Kotliansky V.,  
 RA Gotwals P.J., Karpusas M.;  
 RT "Crystal structure of the alphabeta1 integrin I-domain: insights into  
 integrin I-domain function."  
 RL FEBS Lett. 452:379-385(1999).  
 CC -I- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND  
 COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-  
 E-R IN COLLAGEN.  
 CC -I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1  
 ASSOCIATES WITH BETA-1.  
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -I- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -I- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -I- SIMILARITY: Contains 1 VWFA domain.  
 CC -I- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 CC -----

```

DR EMBL; X52140; CAA36384.1; -.
DR PIR; A35854; A35854.
DR PDB; 1CK4; O3-MAY-00.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; integrin_A; 1.
DR Pfam; PF00092; vwa; 1.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02334; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Calcium; Magnesium; 3D-structure.
FT SIGNAL 1 28
FT CHAIN 29 1180 INTEGRIN_ALPHA-1.
FT DOMAIN 29 1142 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1143 1165 POTENTIAL.
FT DOMAIN 1166 1180 CYTOPLASMIC (POTENTIAL).
FT REPEAT 44 103 FG-GAP 1.
FT REPEAT 2 2 FG-GAP 2.
FT DOMAIN 175 388 VWFA.
FT REPEAT 377 432 FG-GAP 3.
FT REPEAT 433 484 FG-GAP 4.
FT REPEAT 485 565 FG-GAP 5.
FT REPEAT 567 626 FG-GAP 6.
FT REPEAT 629 681 FG-GAP 7.
FT CA_BIND 497 505 POTENTIAL.
FT CA_BIND 579 587 POTENTIAL.
FT CA_BIND 641 649 POTENTIAL.
FT SITE 1168 1172 GFPR MOTIF.
FT DISULFID 82 92 BY SIMILARITY.
FT DISULFID 687 696 BY SIMILARITY.
FT DISULFID 702 755 BY SIMILARITY.
FT DISULFID 807 813 BY SIMILARITY.
FT DISULFID 877 885 BY SIMILARITY.
FT DISULFID 1029 1082 BY SIMILARITY.
FT DISULFID 1066 1073 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 317 317 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 698 698 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 747 779 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 820 820 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 882 882 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 907 907 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 965 965 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 973 973 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1007 1007 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1084 1084 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1103 1103 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1114 1114 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SEQUENCE 1180 AA; 130808 MW; 8B5DA2B802362EB4 CRC64;

```

Query Match 1.3%; Score 15; DB 1; Length 1180;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 164 DIVIVLDGNSIYPW 178
DB 172 DIVIVLDGNSIYPW 186

```

```

RESULT 5
ITAG HUMAN
ID ITAG HUMAN STANDARD; PRT; 1167 AA.
AC O75578; Q9UHZ8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-10 precursor.
GN ITGA10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
EN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Articular chondrocytes;
RX MEDLINE=98352078; PubMed=9685391;
RA Camper L., Hellman U., Lundgren-Akerlund E.;
RT "Isolation, cloning, and sequence analysis of the integrin subunit
alpha10, a beta1-associated collagen binding integrin expressed on
chondrocytes.";
RT J. Biol. Chem. 273:20383-20389(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Endothelial cells, and Heart;
RX MEDLINE=20169197; PubMed=10702680;
RA Lehnert K., Ni J., Leung Z., Gough S.M., Morris C.M., Liu D.,
Wang S.-X., Langley R., Krissansen G.W.;
RT "The integrin alpha10 subunit: expression pattern, partial gene
structure, and chromosomal localization.";
RT Cytogenet. Cell Genet. 87:238-244(1999).
CC -!- FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Widely expressed with highest expression in
muscle and heart. Found in articular cartilage.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF074015; AAC31952.1; -.
DR EMBL; AF112345; AAF21944.1; -.
DR EMBL; AF172723; AAF61638.1; -.
DR HSSP; P17301; LAOX.
DR Genew; HGNC:6135; ITGA10.
DR MIM; 604042; -.
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO; GO:0005518; F:collagen binding; TAS.
DR GO; GO:0007160; P:cell-matrix adhesion; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.
DR PROSITE; PS02334; VWFA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Signal; Repeat; Calcium; Magnesium.

```

FT SIGNAL 1 22  
 FT CHAIN 23 1167  
 FT DOMAIN 23 1122  
 FT TRANSMEM 1123 1145  
 FT DOMAIN 1146 1167  
 FT REPEAT 38 97  
 FT REPEAT 7 7  
 FT DOMAIN 167 350  
 FT REPEAT 365 427  
 FT REPEAT 428 482  
 FT REPEAT 546 605  
 FT REPEAT 608 660  
 FT DOMAIN 1134 1140  
 FT CA\_BIND 494 502  
 FT CA\_BIND 558 566  
 FT CA\_BIND 620 628  
 FT DISULFID 76 86  
 FT DISULFID 666 675  
 FT DISULFID 681 736  
 FT DISULFID 789 795  
 FT CARBOHYD 98 98  
 FT CARBOHYD 234 234  
 FT CARBOHYD 336 336  
 FT CARBOHYD 364 364  
 FT CARBOHYD 733 733  
 FT CARBOHYD 763 763  
 FT CARBOHYD 839 839  
 FT CARBOHYD 921 921  
 FT CARBOHYD 1011 1011  
 FT CARBOHYD 1018 1018  
 FT CARBOHYD 1039 1039  
 FT CONFLICT 844 844  
 FT CONFLICT 909 909  
 FT CONFLICT 926 926  
 SQ SEQUENCE 1167 AA; 127573 MW; A57D3A1C25C1AAE0 CRC64;

Query Match 1.1%; Score 13; DB 1; Length 1167;  
 Best Local Similarity 100.0%; Pred. No. 0.00054;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 VIVLDGNSNIYPW 178  
 DB 169 VIVLDGNSNIYPW 181  
 |||||

RESULT 6  
 ID ITA2 BOVIN STANDARD; PRT; 1170 AA.  
 AC P53710;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DE 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GPIa)  
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (Fragment).  
 GN ITGA2.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94193647; PubMed=75111592;  
 RA Kamata T., Puzon W., Takada Y.;  
 RT "Identification of putative ligand binding sites within I domain of  
 integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";  
 RL J. Biol. Chem. 269:19659-19663(1994).  
 CC FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,  
 CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND B-CADHERIN. IT  
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-R IN  
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER  
 CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE

CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
 CC EXTRACELLULAR MATRIX.  
 CC -!- SUBUNIT; HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2  
 CC -!- ASSOCIATES WITH BETA-1.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; L25886; AAB59255.1; -;  
 DR PIR; I45914; I45914.  
 DR HSSP; P17301; LAOX.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VMF\_A.  
 DR Pfam; PF01839; FG-GAP; 3\_A.  
 DR Pfam; PF00357; Integrin\_A; 1.  
 DR Pfam; PF00092; vwa; 1.  
 DR SMART; SM00191; Int\_alpha; 5.  
 DR SMART; SM00327; VWFA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS50234; VWFA; 1.  
 DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 18  
 FT CHAIN 19 1170  
 FT DOMAIN 19 1170  
 FT TRANSMEM 1122 1143  
 FT DOMAIN 1144 1170  
 FT REPEAT 34 92  
 FT REPEAT 7 7  
 FT DOMAIN 177 367  
 FT REPEAT 423 475  
 FT REPEAT 477 538  
 FT REPEAT 540 599  
 FT REPEAT 604 656  
 FT CA\_BIND 488 496  
 FT CA\_BIND 552 560  
 FT CA\_BIND 616 624  
 FT SITE 472 474  
 FT SITE 1146 1150  
 FT DISULFID 72 81  
 FT DISULFID 669 726  
 FT DISULFID 778 784  
 FT DISULFID 854 865  
 FT DISULFID 1008 1039  
 FT DISULFID 1044 1049  
 FT CARBOHYD 94 94  
 FT CARBOHYD 101 101  
 FT CARBOHYD 332 332  
 FT CARBOHYD 421 421  
 FT CARBOHYD 449 449  
 FT CARBOHYD 464 464  
 FT CARBOHYD 688 688  
 FT CARBOHYD 748 748  
 FT CARBOHYD 945 945  
 FT CARBOHYD 1063 1063  
 FT CARBOHYD 1070 1070  
 FT VARIANT 580 580  
 FT VARIANT 588 588  
 FT VARIANT 725 725  
 FT SEQUENCE 1170 AA; 128929 MW; BECEFF1C5F2448FB1 CRC64;

Query Match 0.9%; Score 11; DB 1; Length 1170;  
 Best Local Similarity 100.0%; Pred. No. 0.05;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAPMY 505  
 |||||  
 DB 495 TDVLLVGAPMY 505

RESULT 7

IT2A2 HUMAN STANDARD; PRT; 1181 AA.

AC P17301;  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GPIa)  
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).  
 GN ITGA2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-44.  
 RC TISSUE-Endothelial cells;  
 RX MEDLINE=89308879; PubMed=2545729;  
 RA Takada Y., Hemler M.E.;  
 RT "The primary structure of the VLA-2/collagen receptor alpha 2 subunit  
 (platelet GPIa): homology to other integrins and the presence of a  
 RT possible collagen-binding domain.";  
 RL J. Cell Biol. 109:397-407(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Rieder M.J., Armet T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,  
 RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 168-368.  
 RX MEDLINE=98019223; PubMed=9353312;  
 RA Emley J., King S.L., Bergelson J.M., Liddington R.C.;  
 RT "Crystal structure of the I domain from integrin alpha2beta1.";  
 RL J. Biol. Chem. 272:28512-28517(1997).  
 RN [4]  
 RP VARIANT HPA-5 (BR).  
 RX MEDLINE=94043762; PubMed=7901236;  
 RA Santoso S., Kalb R., Waika M., Kiefel V., Mueller-Eckhardt C.,  
 RA Newman P.J.;  
 RT "The human platelet alloantigens Br(a) and Br(b) are associated with a  
 RT single amino acid polymorphism on glycoprotein Ia (integrin subunit  
 RT alpha 2).";  
 RL J. Clin. Invest. 92:2427-2432(1993).  
 RN [5]  
 RP VARIANT GLU-534.  
 RX MEDLINE=20206009; PubMed=10741412;  
 RA Kroll H., Gardemann A., Fechter A., Haberbesch W., Santoso S.;  
 RT "The impact of the glycoprotein Ia collagen receptor subunit A1648G  
 RT gene polymorphism on coronary artery disease and acute myocardial  
 RT infarction.";  
 RL Thromb. Haemost. 83:392-395(2000).  
 CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,  
 CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT  
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-E-R IN  
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER  
 CC CELLS TO COLLAGENS. MODULATION OF COLLAGEN AND COLLAGENASE GENE  
 CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
 CC EXTRACELLULAR MATRIX.  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-2  
 CC associates with beta-1. Interacts with HPSS.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.

-!- POLYMORPHISM: Position 534 is associated with platelet-specific  
 alloantigen HPA-5 (BR). HPA-5A/BR(A) has Lys-534 and HPA-5B/BR(B)  
 has Glu-534. HPA-5B is involved in neonatal alloimmune  
 thrombocytopenia (NAIT or NATP). The K534E polymorphism may play a  
 role in coronary artery disease (CAD).  
 -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 -!- SIMILARITY: Contains 1 VWFA domain.  
 -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 -!- DATABASE: NAME=PROV; NOTE=CD guide CD49b entry;  
 WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd49b.htm".

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

EMBL; X17033; CAA34894.1; --  
 EMBL; AFS12556; AAM34795.1; --  
 PIR; A33998; A33998.  
 PDB; 1AOX; 25-NOV-98.  
 PDB; 1DZI; 02-AUG-01.  
 Genew; HGNC:6137; ITGA2.  
 MIM; 192374; --  
 DR GO; 0008305; C:integrin complex; TAS.  
 DR GO; 0005886; C:plasma membrane; TAS.  
 DR GO; 0004895; F:cell adhesion receptor activity; TAS.  
 DR GO; 0005518; F:collagen binding; TAS.  
 DR GO; 0007596; P:blood coagulation; TAS.  
 DR GO; 0007160; P:cell-matrix adhesion; TAS.  
 DR GO; 0007397; P:histogenesis and organogenesis; TAS.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWFA.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00357; integrin\_A; 1.  
 DR Pfam; PF00092; vwa; 1.  
 DR SMART; SM00191; Int\_alpha; 5.  
 DR SMART; SM00327; VWFA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS50234; VWFA; 1.  
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Platelet; Signal; Repeat; Calcium; Magnesium; Polymorphism;  
 3D-structure.

FT SIGNAL	1	29	INTEGRIN ALPHA-2
FT CHAIN	30	1181	EXTRACELLULAR (POTENTIAL).
FT DOMAIN	30	1132	POTENTIAL.
FT TRANSMEM	1133	1154	CYTOPLASMIC (POTENTIAL).
FT DOMAIN	1155	1181	INTERACTION WITH HPSS.
FT REPEAT	45	103	FG-GAP 1.
FT REPEAT	?	?	FG-GAP 2.
FT DOMAIN	188	378	VWFA.
FT REPEAT	378	433	FG-GAP 3.
FT REPEAT	434	486	FG-GAP 4.
FT REPEAT	488	549	FG-GAP 5.
FT REPEAT	551	610	FG-GAP 6.
FT REPEAT	615	667	FG-GAP 7.
FT CA_BIND	499	507	POTENTIAL.
FT CA_BIND	563	571	POTENTIAL.
FT CA_BIND	627	635	POTENTIAL.
FT SITE	1157	1161	GPFR MOTIF.
FT DISULFID	83	92	BY SIMILARITY.
FT DISULFID	680	737	BY SIMILARITY.
FT DISULFID	789	795	BY SIMILARITY.
FT DISULFID	865	876	BY SIMILARITY.
FT DISULFID	1019	1050	BY SIMILARITY.
FT DISULFID	1055	1060	BY SIMILARITY.
FT CARBOHYD	105	105	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	112	112	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	343	343	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	432	432	N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 693 693 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1057 1057 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 1081 1081 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT VARIANT 534 534 K -> E (IN ALLOANTIGEN HPA-59;
  GbSNP:1801106).
  /FTid=VAR_003977.
FT TURN 170 171
FT STRAND 173 180
FT TURN 183 184
FT TURN 188 199
FT TURN 200 201
FT STRAND 204 204
FT TURN 206 207
FT STRAND 209 216
FT STRAND 220 224
FT TURN 226 228
FT TURN 232 240
FT TURN 241 241
FT TURN 252 262
FT TURN 263 264
FT TURN 265 268
FT TURN 269 269
FT STRAND 275 282
FT TURN 289 291
FT TURN 292 301
FT TURN 302 303
FT STRAND 304 311
FT TURN 313 317
FT TURN 318 319
FT TURN 323 330
FT TURN 331 332
FT TURN 337 340
FT STRAND 341 344
FT TURN 347 353
FT TURN 354 362
FT TURN 363 363
SQ SEQUENCE 1181 AA, 129295 MW; 7B1B7ED968A94070 CRC64;

Query Match 0.9%; Score 11; DB 1; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAPMY 505
Db 506 TDVLLVGAPMY 516
|||||

RESULT 8
CAH4_BOVIN STANDARD; PRT; 312 AA.
AC Q95323;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Carbonic anhydrase IV precursor (EC 4.2.1.1) (Carbonate dehydratase IV) (CA-IV).
GN CA4.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Holstein; TISSUE=Kidney;
RA Tamai S.;
CC Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Reversible hydration of carbon dioxide.
CC -!- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
CC -!- COFACTOR: Zinc.

```

```

CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U58870; AAB09466.1; -.
CC HSSP; P22748; 1ZNC.
DR InterPro; IPR001148; Euk_Coanhd.
DR Pfam; PF00194; carb anhydrase; 1.
DR ProDom; PD000865; Euk_Coanhd; 1.
DR PROSITE; PS00162; EUK_CO2_ANGHYDRASE; FALSE NEG.
KW Glycoprotein; Lipoprotein; GPI-anchor; Membrane; Lyase; Zinc; Signal.
FT SIGNAL 1 18
FT CHAIN 19 284
FT PROPEP 285 312
FT METAL 115 115
FT METAL 117 117
FT METAL 140 140
FT DISULFID 24 36
FT DISULFID 46 229
FT LIPID 284 284
FT SEQUENCE 312 AA; 35151 MW; BAA320C09426351 CRC64;

Query Match 0.8%; Score 9; DB 1; Length 312;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVLA 1162
Db 3 LLLALLVLA 11
|||||

RESULT 9
LEU3_SCHPO STANDARD; PRT; 371 AA.
AC P18869;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 3-isopropylmalate dehydrogenase (EC 1.1.1.85) (Beta-IPM dehydrogenase)
DE (IMDH) (3-IPM-DH).
GN LEU1 OR SPEC1A4.02C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89106270; PubMed=3063400;
RA Kikuchi Y., Kitazawa Y., Shimatake G., Yamamoto M.;
RT "The primary structure of the leu1+ gene of Schizosaccharomycetes
RT pombe".
RL Curr. Genet. 14:375-379(1988).
RC [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Gobie A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

```



RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volktaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Woestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallard J., Talleda V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.K., Paulsen O., Potashkin J.,  
 RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.,  
 RT The genome sequence of *Schizosaccharomyces pombe*.  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: Catalyzes the oxidation of 3-carboxy-2-hydroxy-4-  
 CC methylpentanoate (3-isopropylmalate) to 3-carboxy-4-methyl-2-  
 CC oxopentanoate. The product decarboxylates to 4-methyl-2-  
 CC oxopentanoate.  
 CC -!- CATALYTIC ACTIVITY: 3-carboxy-2-hydroxy-4-methylpentanoate +  
 CC NAD(+) = 3-carboxy-4-methyl-2-oxopentanoate + NADH.  
 CC -!- PATHWAY: Leucine biosynthesis; third step.  
 CC -!- SUBUNIT: Homodimer (by similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to the isocitrate and isopropylmalate  
 CC dehydrogenases family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M36910; AAA35316.1; --  
 CC EMBL; AL031174; CAA20106.1; --  
 CC PIR; T43407; T43407.  
 CC HSSP; P12010; 2AVO.  
 CC GeneDB\_Spombe; SPBC1A4.02c; --  
 CC InterPro; IPR001804; IsoDH.  
 CC InterPro; IPR004429; LeuB.  
 CC Pfam; PF00180; IsoDH; 1.  
 CC TIGRFAMs; TIGR00169; leuB; 1.  
 CC PROSITE; PS00470; IDH IMDH; 1.  
 CC Oxidoreductase; Leucine biosynthesis; NAD.  
 KW SEQUENCE 371 AA; 39732 MW; 65AA2B6AA94D45EE CRC64;  
 QY Query Match 0.8%; Score 9; DB 1; Length 371;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 378 DGVLLGAVG 386  
 69 DGVLLGAVG 77  
 RESULT 10  
 ID ITA2 MOUSE  
 AC Q62469; Q62463; PRT; 1178 AA.  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GPIa)  
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).  
 GN ITGA2.  
 GN Mus musculus (Mouse).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=CS7BL/6 X CBA; TISSUE=Lung;  
 MDLINE=94363406; PubMed=8081889;  
 RA Edelman J.M., Chan B.M., Unival S., Onodera H., Wang D.Z.,  
 RA Damjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.,  
 RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but  
 RT not virus binding."  
 RL Cell Adhes. Commun. 2:131-143(1994).  
 [2]  
 RC SEQUENCE OF 450-1178 FROM N.A.  
 RC TISSUE=Lung;  
 MDLINE=94355691; PubMed=7521231;  
 RA Wu J.B., Santoro S.A.;  
 RT "Complex patterns of expression suggest extensive roles for the alpha  
 RT 2 beta 1 integrin in murine development."  
 RL Dev. Dyn. 199:292-314(1994).  
 CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING  
 CC RESPONSIBLE FOR ADHESION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,  
 CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED  
 CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN  
 CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION  
 CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.  
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2  
 CC ASSOCIATES WITH BETA-1.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Z29987; CAA82877.1; --  
 CC EMBL; X75427; CAA53178.1; --  
 CC PIR; S44142; S44142.  
 CC HSSP; P17301; IAOX.  
 CC MGD; MGI:96600; Itga2.  
 CC InterPro; IPR000413; Integrin\_alpha.  
 CC InterPro; IPR002035; VWF\_A.  
 CC Pfam; PF01839; FG-GAP; 3\_  
 CC Pfam; PF00357; integrin\_A; 1.  
 CC SMART; SM00191; Int\_alpha; 4.  
 CC SMART; SM00327; VWFA; 1.  
 CC PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 CC PROSITE; PS00234; VWFA; 1.  
 CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Platelet; Signal; Repeat; Calcium; Magnesium.  
 FT SIGNAL 1 26  
 FT CHAIN 27 1178  
 FT DOMAIN 27 1129  
 FT TRANSMEM 1130 1151  
 FT DOMAIN 1152 1178  
 FT REPEAT 42 100  
 FT REPEAT ? ?  
 FT REPEAT 185 375  
 FT DOMAIN ? ?  
 FT REPEAT 431 483  
 FT REPEAT 485 546  
 FT REPEAT 548 607  
 FT REPEAT 612 664  
 FT CA\_BIND 496 504  
 FT POTENTIAL.  
 FT CYTOPLASMIC (POTENTIAL).  
 FT EXTRACELLULAR (POTENTIAL).  
 FT INTEGRIN ALPHA-2.  
 FT BY SIMILARITY.

```

FT CA BIND 560 568 POTENTIAL.
FT CA_BIND 624 632 POTENTIAL.
FT SITE 480 482 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 1154 1158 GPCR MOTIF.
FT DISULFID 80 89 BY SIMILARITY.
FT DISULFID 677 734 BY SIMILARITY.
FT DISULFID 786 792 BY SIMILARITY.
FT DISULFID 862 873 BY SIMILARITY.
FT DISULFID 1016 1047 BY SIMILARITY.
FT DISULFID 1052 1057 BY SIMILARITY.
FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 429 429 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 472 472 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 696 696 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1054 1054 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1178 AA; 128926 MW; 1F194B9C0240F465 CRC64;

Query Match 0.8%; Score 9; DB 1; Length 1178;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVCGAP 503
DB 503 TDVLLVCGAP 511
|||||

RESULT 11
CD24_MOUSE
ID CD24_MOUSE STANDARD; PRT; 76 AA.
AC P24807; P26691;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Signal transducer CD24 precursor (M1/69-J11D heat stable antigen)
DE (HSA) (Nectadrin) (LY-52) (X82 heat stable antigen) (R13-AG).
GN CD24 OR CD24A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
(1)
RN SEQUENCE FROM N.A.
RX MEDLINE=90361906; PubMed=2118158;
RA Kay R., Takei F., Humphries R.K.;
RT "Expression cloning of a cDNA encoding M1/69-J11D heat-stable
antigens."
RL J. Immunol. 145:1952-1959(1990).
(2)
RN SEQUENCE FROM N.A.
RX STRAIN=C57BL/6 X CBA, and Swiss albino X BALB/c; TISSUE=Spleen;
RA MEDLINE=91209380; PubMed=2019286;
RA Wenger R.H., Ayane M., Bose R., Koehler G., Nielsen P.J.;
RT "The genes for a mouse hematopoietic differentiation marker called
the heat-stable antigen."
RL Eur. J. Immunol. 21:1039-1046(1991).
(3)
RN REVISIONS.
RA Nielsen P.J.;
RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
(4)
RN SEQUENCE FROM N.A.
RX STRAIN=C57BL/6 X CBA; TISSUE=Spleen;
RA MEDLINE=94043127; PubMed=8226859;
RA Wenger R.H., Rochelle J.M., Seldin M.F., Koehler G., Nielsen P.J.;
RT "The heat stable antigen (mouse CD24) gene is differentially
regulated but has a housekeeping promoter."
RL J. Biol. Chem. 268:23345-23352(1993).
(5)
RN SEQUENCE OF 27-53.

```

```

RC STRAIN=C57BL/6;
RX MEDLINE=92412120; PubMed=1530634;
RA Hitsumoto Y., Nakano A., Ohnishi H., Hamada F., Sabeki S.,
RA Takeuchi N.;
RT "Purification of the murine heat-stable antigen from erythrocytes.";
RL Biochem. Biophys. Res. Commun. 187:773-777(1992).
CC -!- FUNCTION: May have a specific role to play in early thymocyte
development.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- TISSUE SPECIFICITY: In lymphoid, myeloid, and erythroid cells.
CC -!- PTM: Extensively O-glycosylated (By similarity).
CC -!- SIMILARITY: TO OTHER MAMMALIAN SPECIES CD24.
-----
This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; M58661; AAA39481.1; -
DR EMBL; X56469; CAA39841.1; -
DR EMBL; X72910; CAA51415.1; -
DR EMBL; X53825; CAA37822.1; -
DR PIR; A43537; A43537
DR MGd; MG1:88323; Cd24a.
KW Glycoprotein; GPI-anchor; Membrane; Signal; Antigen; Lipoprotein.
FT SIGNAL 1 26
FT CHAIN 27 53 SIGNAL TRANSDUCER CD24.
FT PROPEP 54 76 REMOVED IN MATURE FORM (POTENTIAL).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .).
FT CARBOHYD 30 30 O-LINKED (PROBABLE).
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .).
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .).
FT LIPID 53 53 GPI-anchor amidated glycine (Potential).
SQ SEQUENCE 76 AA; 7797 MW; 6853F12B33625EB CRC64;

Query Match 0.7%; Score 8; DB 1; Length 76;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 12 GLLLLALL 19
|||||

RESULT 12
CD24_RAT
ID CD24_RAT STANDARD; PRT; 76 AA.
AC Q07490;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Signal transducer CD24 precursor (Heat stable antigen) (HSA)
DE (Nectadrin).
GN CD24A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
(1)
RN SEQUENCE FROM N.A.
RX STRAIN=Wistar; TISSUE=Embryonic brain;
RX MEDLINE=94122434; PubMed=8292828;
RA Shirasawa T., Akashi T., Sakamoto K., Takahashi H., Maruyama N.,
RA Hirokawa K.;
RT "Gene expression of CD24 core peptide molecule in developing brain
and developing non-neural tissues."
RL Dev. Dyn. 198:1-13(1993).
(2)
RN SEQUENCE FROM N.A.
RX STRAIN=Fischer;

```

RX MEDLINE=97157759; PubMed=9004038;  
RA Magaldi T.A., Barrandon Y.;  
RT "CD24 (heat stable antigen, nectadrin), a novel keratinocyte  
RT differentiation marker, is preferentially expressed in areas of the  
RT hair follicle containing the colony-forming cells.";  
RL J. Cell Sci. 109:3035-3045(1998).  
CC -!- FUNCTION: May have a pivotal role in cell differentiation. The  
CC triggering mechanism of signal transduction may be due to the  
CC interactions of differentiating cells with the matrix substrate  
CC via the carbohydrate structure of the molecule. In this way, the  
CC signal transducer can play very different roles in different cell  
CC types as a direct consequence of its glycosylation.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
CC -!- TISSUE SPECIFICITY: Expressed in the central nervous system, in  
CC postmitotic cells of spinal cord, hindbrain, midbrain and  
CC forebrain. Expressed in epithelium during the development of non-  
CC neural tissues. Expressed in tooth development, specifically in  
CC mesenchymal cells differentiating into odontoblast in dental  
CC papilla, as well as in the developing eye and hair follicle.  
CC -!- DEVELOPMENTAL STAGE: Detected in primitive ectoderm, mesoderm and  
CC ventral endoderm; down-regulated when organogenesis is completed.  
CC -!- PTM: Extensively O-glycosylated (by similarity). The carbohydrate  
CC structure may be regulated in a tissue-specific and developmental  
CC stage-specific manner.  
CC -!- SIMILARITY: TO OTHER MAMMALIAN SPECIES CD24.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
DR EMBL; Z11663; CAA77731.1; -;  
DR EMBL; U49062; AAA91470.1; -;  
DR PIR; I53107; I53107.  
KW Glycoprotein; GPI-anchor; Membrane; Signal; Differentiation;  
KW Lipoprotein.  
FT SIGNAL 1 26 POTENTIAL.  
FT CHAIN 27 56 SIGNAL TRANSDUCER CD24. (BY SIMILARITY).  
FT PROPEP 57 76 REMOVED IN NATURE FORM (POTENTIAL).  
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT LIPID 56 56 GPI-anchor amidated serine (Potential).  
FT SEQUENCE 76 AA; 7862 MW; 4284670BC39D958 CRC64;  
Query Match 0.7%; Score 8; DB 1; Length 76;  
Best Local Similarity 100.0%; Pred. No. 4.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1152 GULLALL 1159  
Db 12 GULLALL 19  
RESULT 13  
ID CD24\_HUMAN STANDARD; PRT; 80 AA.  
AC P25063; Q16257;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Signal transducer CD24 precursor.  
GN CD24 OR CD24A.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91332458; PubMed=1831224;  
RA Kay R., Rosten P.M., Humphries R.K.;  
RT "CD24, a signal transducer modulating B cell activation responses, is  
RT a very short peptide with a glycosyl phosphatidylinositol membrane  
RT anchor.";  
RL J. Immunol. 147:1412-1416(1991).  
RN [2]  
RP SEQUENCE FROM N.A., AND VARIANT VAL-57.  
RX MEDLINE=93007871; PubMed=1327504;  
RA Jackson D., Waibel R., Weber E., Bell J., Stabel R.A.;  
RT "CD24, a signal-transducing molecule expressed on human B cells, is a  
RT major surface antigen on small cell lung carcinomas.";  
RL Cancer Res. 52:5264-5270(1992).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feirgold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins P.S., Wagner L., Shenman C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raba S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.M.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP SEQUENCE OF 1-76 FROM N.A.  
RX MEDLINE=95048364; PubMed=7959762;  
RA Hough M.R., Rosten P.M., Sexton T.L., Kay R., Humphries R.K.;  
RT "Mapping of CD24 and homologous sequences to multiple chromosomal  
RT loci.";  
RL Genomics 22:154-161(1994).  
CC -!- FUNCTION: Modulates B-cell activation responses. Signaling could  
CC be triggered by the binding of a lectin-like ligand to the CD24  
CC carbohydrates, and transduced by the release of second messengers  
CC derived from the GPI-anchor. Promotes AG-dependent proliferation  
CC of B-cells, and prevents their terminal differentiation into  
CC antibody-forming cells.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
CC -!- TISSUE SPECIFICITY: B-cells.  
CC -!- INDUCTION: Expression is lost when primary B-cells are induced to  
CC differentiate in antibody-forming cells.  
CC -!- PTM: Extensively O-glycosylated.  
CC -!- SIMILARITY: TO OTHER MAMMALIAN SPECIES CD24.  
CC -!- DATABASE: NAME=PRO; NOT=CD guide CD24 entry;  
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd24.htm".  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
DR EMBL; M58664; AAA35665.1; -;  
DR EMBL; X69397; CAA49195.1; -;  
DR EMBL; L33930; AAB58307.1; -;  
DR EMBL; BC007674; AAB07674.1; -;  
DR EMBL; S75311; AAD14170.1; ALT\_INIT.  
DR PIR; I56114; A48996.  
DR Genew; HGNC:1645; CD24.

DR MM; 600074; --  
DR GO; GO:0006959; P:humoral immune response; TAS.  
KW Glycoprotein; GPI-anchor; Membrane; Signal; Polymorphism; Lipoprotein.  
FT SIGNAL 1 26 POTENTIAL.  
FT CHAIN 27 59 SIGNAL TRANSDUCER CD24.  
FT PROPEP 60 80 REMOVED IN MATURE FORM (POTENTIAL).  
FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT LIPID 59 59 GPI-anchor amidated glycine (Potential).  
FT VARIANT 57 57 A -> V (in dbSNP:8734).  
FT FTID=VAR 016156.  
FT CONFLICT 12 12 G -> W (IN REF. 4).  
FT CONFLICT 44 44 S -> T (IN REF. 4).  
SQ SEQUENCE 80 AA; 8083 MW; DB1988B6808F833F CRC64;  
  
Query Match 0.7%; Score 8; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 5.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1152 GLLLLALL 1159  
Db 12 GLLLLALL 19  
  
RESULT 14  
LST1 HUMAN STANDARD; PRT; 97 AA.  
AC Q00453; Q00452; Q00454; Q13669; Q9UUR5; Q9UUR6; Q9UUR7; Q9UUR8;  
AC Q9UUR1;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Leukocyte specific transcript 1 protein (B144 protein).  
GN LST1 OR B144.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 3), TISSUE SPECIFICITY, AND INDUCTION.  
RC TISSUE=Blood;  
RX MEDLINE=9600655; PubMed=7590964;  
RA Holzinger I., de Baey A., Messer G., Kick G., Zwiarsina H.,  
RA Weiss E.H.;  
RT "Cloning and genomic characterization of LST1: a new gene in the human  
TNF region.";  
RL Immunogenetics 42:315-322(1995).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 4), TISSUE SPECIFICITY, AND  
INDUCTION.  
RC TISSUE=Lymphoblast;  
RX MEDLINE=38035883; PubMed=9367684;  
RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,  
RA Weiss E.H.;  
RT "Complex expression pattern of the TNF region gene LST1 through  
differential regulation, initiation, and alternative splicing.";  
RL Genomics 45:591-600(1997).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORMS 5; 6; 7; 8 AND 9), FUNCTION, AND  
SUBCELLULAR LOCATION.  
RC TISSUE=Periphereal blood;  
RX MEDLINE=20171517; PubMed=10706707;  
RA Rollinger-Holzinger I., Bibl B., Pauly M., Grieser U., Hentges F.,  
RA Auer B., Fall G., Schratzberger P., Niederwieser D., Weiss E.H.,  
RA Zwiarsina H.;  
RT "LST1: a gene with extensive alternative splicing and immunomodulatory  
function.";  
RL J. Immunol. 164:3169-3176(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,  
RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,  
RA Lasky S., Hood L.;

RT "Sequence of the human major histocompatibility complex class III  
region.";  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Hirakawa M., Yamaguchi H., Imai K., Shimada J.,  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP FUNCTION, AND SUBCELLULAR LOCATION.  
RX MEDLINE=21372017; PubMed=11478849;  
RA Raghunathan A., Sivakamasundari R., Wolenski J., Poddar R.,  
RA Weissman S.M.;  
RT "Functional analysis of B144/LST1: a gene in the tumor necrosis  
factor cluster that induces formation of long filopodia in eukaryotic  
cells.";  
RL Exp. Cell Res. 268:230-244(2001).  
CC -!- FUNCTION: Possible role in modulating immune responses. Isoforms 1  
and 2 have an inhibitory effect on lymphocyte proliferation.  
Induces morphological changes including production of filopodia  
and microspikes when overexpressed in a variety of cell types and  
may be involved in dendritic cell maturation.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Also detected in  
a perinuclear region corresponding to the localization of the  
Golgi apparatus and throughout the cytoplasm.  
CC -!- ALTERNATIVE PRODUCTS:  
Event-Alternative splicing; Named isoforms=9;  
Comment-Additional isoforms seem to exist;  
Name=1; Synonyms=LST1/A;  
IsoId=O00453-1; Sequence=Displayed;  
Name=2; Synonyms=LST1/C;  
IsoId=O00453-2; Sequence=VSP\_050578;  
Name=3; Synonyms=PLST1;  
IsoId=O00453-3; Sequence=VSP\_050579;  
Name=4; Synonyms=LST1/E;  
IsoId=O00453-4; Sequence=VSP\_050577;  
Name=5; Synonyms=LST1/K;  
IsoId=O00453-5; Sequence=VSP\_050584;  
Name=6; Synonyms=LST1/L;  
IsoId=O00453-6; Sequence=VSP\_050583; VSP\_050587;  
Name=7; Synonyms=LST1/J;  
IsoId=O00453-7; Sequence=VSP\_050586;  
Name=8; Synonyms=LST1/M;  
IsoId=O00453-8; Sequence=VSP\_050580;  
Name=9; Synonyms=LST1/N;  
IsoId=O00453-9; Sequence=VSP\_050581; VSP\_050582;  
CC -!- TISSUE SPECIFICITY: Expressed in lung, tonsil, thymus, placenta,  
kidney, fetal spleen, fetal liver and brain.  
CC -!- INDUCTION: By interferon gamma.  
CC -!- SIMILARITY: Belongs to the LST1 family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL; U00921; AAB57724.1; --  
DR EMBL; AF000424; AAB86998.1; --  
DR EMBL; AF000425; AAB86999.1; --  
DR EMBL; AF000426; AAB87000.1; --  
DR EMBL; Y18486; CAB59304.1; --  
DR EMBL; Y18487; CAB59305.1; --  
DR EMBL; Y18488; CAB60038.1; --  
DR EMBL; Y18489; CAB59306.1; --  
DR EMBL; Y18490; CAB59303.1; --  
DR EMBL; AF129756; AAD18090.1; --  
DR EMBL; AP000505; BAB61394.1; --  
DR Gene; HGNC:14189; LST1.  
DR MIM; 109170; --  
DR GO; GO:0006955; P:immune response; NAS.  
DR GO; GO:0009653; P:morphogenesis; NAS.

```
DR InterPro: IPR007775; LST1.  
KW Pfam: PF05083; LST1; 1.  
FT Immune response; Cell shape; Transmembrane; Alternative splicing.  
FT TRANSMEM 8 28  
FT VARSPLIC 1 46  
FT MEDLINE=21534947; PubMed=11677608;  
FT PARKHILL J., DUGAN G., JAMES K.D., THOMSON N.R., PICKARD D., WAIN J.,  
FT CHURCHER C., MUNGALL K.L., BENTLEY S.D., HOLDEN M.T.G., SEBAHIA M.,  
FT BAKER S., BASHAM D., BROOKS K., CHILLINGWORTH T., CONNERTON P.,  
FT CRONIN A., DAVIS P., DAVIES R.M., DOWD L., WHITE N., FARRAR J.,  
FT FELTWEIL T., HANLIN N., HAQUE A., HIEN T.T., HOLTOYD S., JAGELS K.,  
FT KROGH A., LAUSEN T.S., LEATHER S., MOULE S., O'GAORA P., PARRY C.,  
FT QUAIL M.A., RUTHERFORD K., SIMMONDS M., SKELTON J., STEVENS K.,  
FT WHITEHEAD S., BARRELL B.G.;  
FT "Complete genome sequence of a multiple drug resistant Salmonella  
FT enterica serovar Typhi CT18.";  
FT Nature 413:848-852(2001).  
FT [3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=S typhi; STRAIN=Ty2 / ATCC 700931;  
RX MEDLINE=22531367; PubMed=12644504;  
RA Deng W., LIU S.-R., PLUNKETT G. III, MAYHEW G.F., ROSE D.J.,  
RA BURLAND V., KODOYIANNI V., SCHWARTZ D.C., BLATTNER P.R.;  
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2  
RT and CT18.";  
RL J. Bacteriol. 185:2330-2337(2003).  
CC -!- FUNCTION: Required for the cell division process (By similarity).  
CC -!- SUBUNIT: May interact with ftsL (By similarity). Inner membrane  
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane  
CC (Potential). Colocalizes with ftsL to the division site (By  
CC similarity).  
CC -!- SIMILARITY: Belongs to the ftsB family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; AB008833; AL21811.1; -  
DR EMBL; AL627276; CAD05037.1; -  
DR EMBL; AB016843; AA070389.1; -  
DR STYGENE; SG77777; ftsB.  
DR HAMAP; MF 00599; -; 1.  
DR InterPro: IPR007060; DivIC.  
DR Pfam; PF04977; DivIC; 1.  
KW Cell division; Transmembrane; Inner membrane; Coiled coil;  
KW Complete proteome.  
FT DOMAIN 1 3 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 4 21 POTENTIAL.  
FT DOMAIN 22 103 PERIPLASMIC (POTENTIAL).  
FT DOMAIN 30 74 COILED COIL (POTENTIAL).  
FT SEQUENCE 103 AA; 11575 MW; F3B559AB77E8ACEC CRC64;  
SQ SEQUENCE 103 AA; 11575 MW; F3B559AB77E8ACEC CRC64;  
Query Match 0.7%; Score 8; DB 1; Length 103;  
Best Local Similarity 100.0%; Pred. No. 6.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1153 LLLALLLV 1160  
DB 6 LLLALLLV 13  
RESULT 16  
FTSB KLEAE STANDARD; PRT; 105 AA.  
ID _FTSB KLEAE  
AC 09EVY3;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cell division protein ftsB homolog.  
DE "Complete genome sequence of Salmonella enterica serovar Typhimurium  
DE LT2.";  
GN FTSB.  
RP SEQUENCE FROM N.A.  
RC SPECIES=S typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;  
RX MEDLINE=21534948; PubMed=11677609;  
RA McClelland M., SANDERSON K.E., SPIETH J., CLIFTON S.W., LATREILLE P.,  
RA COURTNEY L., FORWOLLIK S., ALI J., DANTE M., DU P., HOU S., LAYMAN D.,  
RA LEONARD S., NGUYEN C., SCOTT K., HOLMES A., GREWAL N., MULVANEY E.,  
RA RYAN E., SUN H., FLOREA L., MILLER W., STONEKING T., NHAN M.,  
RA WATERSTON R., WILSON R.K.;  
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
RT LT2.";  
RP SEQUENCE FROM N.A.  
RC SPECIES=S typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;  
RX MEDLINE=21534948; PubMed=11677609;  
RA McClelland M., SANDERSON K.E., SPIETH J., CLIFTON S.W., LATREILLE P.,  
RA COURTNEY L., FORWOLLIK S., ALI J., DANTE M., DU P., HOU S., LAYMAN D.,  
RA LEONARD S., NGUYEN C., SCOTT K., HOLMES A., GREWAL N., MULVANEY E.,  
RA RYAN E., SUN H., FLOREA L., MILLER W., STONEKING T., NHAN M.,  
RA WATERSTON R., WILSON R.K.;  
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
RT LT2.";  
RP SEQUENCE FROM N.A.  
RC SPECIES=S typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;  
RX MEDLINE=21534948; PubMed=11677609;  
RA McClelland M., SANDERSON K.E., SPIETH J., CLIFTON S.W., LATREILLE P.,  
RA COURTNEY L., FORWOLLIK S., ALI J., DANTE M., DU P., HOU S., LAYMAN D.,  
RA LEONARD S., NGUYEN C., SCOTT K., HOLMES A., GREWAL N., MULVANEY E.,  
RA RYAN E., SUN H., FLOREA L., MILLER W., STONEKING T., NHAN M.,  
RA WATERSTON R., WILSON R.K.;  
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
RT LT2.";
```

OS Klebsiella aerogenes.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Klebsiella.  
 CX NCBI\_TaxID=28451;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=770;  
 RX MEDLINE=20566700; PubMed=11114933;  
 RA Kolk M.M., Kapetanovich L.A., Lawrence J.G.;  
 RT "Alternative pathways for siroheme synthesis in Klebsiella  
 aerogenes";  
 RL J. Bacteriol. 183:328-335(2001).  
 CC -!- FUNCTION: Required for the cell division process (By similarity).  
 CC -!- SUBUNIT: May interact with ftsL (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane  
 CC (potential). Colocalizes with ftsL to the division site (By  
 CC similarity).  
 CC -!- SIMILARITY: Belongs to the ftsB family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR ENBL; AF308468; AAG42461.1; -.  
 DR HAWAP; MF\_00599; -; 1.  
 DR InterPro; IPR007060; DivC.  
 DR Pfam; PF04977; DivC; 1.  
 KW Cell division; Transmembrane; Inner membrane; Coiled coil.  
 FT DOMAIN 1 3 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 4 21 POTENTIAL.  
 FT DOMAIN 22 105 PERIPLASMIC (POTENTIAL).  
 FT DOMAIN 31 74 COILED COIL (POTENTIAL).  
 SQ SEQUENCE 105 AA; 11951 MW; 530471363FD3D112A CRC64;  
 Query Match 0.7%; Score 8; DB 1; Length 105;  
 Best Local Similarity 100.0%; Pred. No. 6.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1153 LLLLLLV 1160  
 DB 6 LLLLLLV 13  
 RESULT 17  
 IL22\_HUMAN  
 ID IL22\_HUMAN STANDARD; PRT; 179 AA.  
 AC Q9GZX6;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Interleukin-22 precursor (IL-22) (IL-10-related T-cell-derived  
 DE inducible factor) (IL-TIP) (UNQ3099/PRO10096).  
 GN IL22 OR IL2IF.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20420346; PubMed=10954742;  
 RA Dumoutier L., Van Roost E., Colau D., Renauld J.-C.;  
 RT "Human interleukin-10-related T cell-derived inducible factor:  
 RT molecular cloning and functional characterization as an hepatocyte-  
 RT stimulating factor";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:10144-10149(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21069354; PubMed=11197690;  
 RA Dumoutier L., Van Roost E., Colau D., Amey G., Michaux L.,

RA Renauld J.-C.;  
 RT "IL-TIP/IL-22: genomic organization and mapping of the human and mouse  
 RT genes";  
 RL Genes Immun. 1:488-494(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20469498; PubMed=10875937;  
 RA Xie M.-H., Aggarwal S., Ho W.-H., Foster J., Zhang Z., Stinson J.,  
 RA Wood W.I., Goddard A.D., Gurney A.L.;  
 RT "Interleukin (IL)-22, a novel human cytokine that signals through the  
 RT interferon receptor-related proteins CRF2-4 and IL-22R";  
 RL J. Biol. Chem. 275:31335-31339(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A., AND VARIANT GLY-158.  
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,  
 RA Nickerson D.A.;  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22887296; PubMed=12975309;  
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,  
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heidens S.,  
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,  
 RA Seshagiri S., Simons L., Singh J., Smith V., Stinson J., Vagts A.,  
 RA Vandlen R., Watanabe C., Wieand D., Woods K., Xie M.-H., Yansura D.,  
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
 RA Godowski P.;  
 RT "The secreted protein discovery initiative (SPDI), a large-scale  
 RT effort to identify novel human secreted and transmembrane proteins: a  
 RT bioinformatics assessment";  
 RL Genome Res. 13:2265-2270(2003).  
 CC -!- FUNCTION: Cytokine that contributes to the inflammatory response  
 CC in vivo.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the IL-10 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ277247; CAC06085.1; -.  
 DR EMBL; AJ277248; CAC19409.1; -.  
 DR EMBL; AF279437; AAG22064.1; -.  
 DR EMBL; AF387519; AAK62468.1; -.  
 DR EMBL; AF388900; AAK89249.1; -.  
 DR Genew; HGNC:14900; IL22.  
 DR MIM; 605330; -.  
 DR GO; GO:0005576; C:extracellular; IC.  
 DR GO; GO:0045518; F:interleukin-22 receptor binding; NAS.  
 DR GO; GO:0006953; P:acute-phase response; NAS.  
 DR GO; GO:0007267; P:cell-cell signaling; IC.  
 DR GO; GO:0006954; P:inflammatory response; NAS.  
 DR InterPro; IPR000098; Interleukin\_10.  
 DR PROSITE; PS00520; INTERLEUKIN\_10; 1.  
 KW Cytokine; Glycoprotein; Signal; Polymorphism.  
 FT SIGNAL 1 33 POTENTIAL.  
 FT CHAIN 34 179 INTERLEUKIN-22.  
 FT CARBOHYD 54 54 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 68 68 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 97 97 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT VARIANT 158 158 S -> G.  
 FT /FTID=VAR\_013078.  
 SQ SEQUENCE 179 AA; 20011 MW; 3C35E64D60CF8767 CRC64;  
 Query Match 0.7%; Score 8; DB 1; Length 179;  
 Best Local Similarity 100.0%; Pred. No. 9.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160  
DB 21 LLLLLLV 28

## RESULT 18

EMBL2\_CAVPO STANDARD; PRT; 234 AA.  
ID EMB2\_CAVPO  
AC P22734;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Eosinophil granule major basic protein 2 precursor (MBP-2).  
GN MBP2.  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
OX NCBI\_TaxID=10141;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Eosinophil;  
RX MEDLINE=91224343; PubMed=2026266;  
RA Aoki I., Shindoh Y., Nishida T., Nakai S., Hong Y.-M., Mio M.,  
RA Saito T., Tasaka K.;  
RT "Comparison of the amino acid and nucleotide sequences between human  
RL FEBS Lett. 282:56-60(1991).  
CC -!- FUNCTION: MBP may play some important roles in the allergic  
CC reactions and inflammations, since MBP is capable of releasing  
CC histamine from mast cells and damaging the epithelial cells of  
CC bronchial tubes. Antiparasitic and antibiotic.  
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific  
CC granule (crystalloid core).  
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D00817; BAA00697.1; -.  
DR PIR; S15102; S15102.  
DR HSSP; P13727; LH8U.  
DR InterPro; IPR002352; Emaior basic.  
DR InterPro; IPR001304; Lectin\_C.  
DR Pfam; PF00059; lectin\_C\_1.  
DR PRINTS; PR00770; EMBJORBASICP.  
DR SMART; SM00034; CLECT; 1.  
DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.  
DR PROSITE; PS00041; C-TYPE LECTIN 2; 1.  
KW Eosinophil; Signal; Immune response; Antibiotic; Lectin;  
KW Multigene family; Glycoprotein.  
FT SIGNAL 1 15 POTENTIAL.  
FT PROPEP 16 115 ACIDIC.  
FT CHAIN 116 234 EOSINOPHIL GRANULE MAJOR BASIC PROTEIN 2.  
FT DOMAIN 133 234 C-TYPE LECTIN (SHORT FORM).  
FT DISULFID 135 232 BY SIMILARITY.  
FT DISULFID 209 224 BY SIMILARITY.  
FT CARBOHYD 69 69 O-LINKED (GLYCOSAMINOGLYCAN) (BY  
FT SIMILARITY)  
SQ SEQUENCE 234 AA; 26140 MW; 7D926A942BF5116F CRC64;

Query Match 0.7%; Score 8; DB 1; Length 234;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160  
DB 4 LLLLLLV 11

## RESULT 19

COMT RAT STANDARD; PRT; 264 AA.  
ID COMT RAT  
AC P22734;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Catechol O-methyltransferase, membrane-bound form (EC 2.1.1.6)  
DE {MB-COMT} [Contains: Catechol O-methyltransferase, soluble form  
DE (S-COMT)].  
GN COMT.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=94107221; PubMed=8280056;  
RA Tenhunen J., Ulmanen I.;  
RT "Production of rat soluble and membrane-bound catechol O-  
RL Biochem. J. 296:595-600(1993).  
CC [2]  
CC SEQUENCE OF 11-264 FROM N.A.  
CC MEDLINE=91033034; PubMed=2227437;  
CC Salminen M., Lundstroem K., Tilgmann C., Savolainen R., Kalkkinen N.,  
CC Ulmanen I.;  
CC "Molecular cloning and characterization of rat liver catechol-O-  
CC methyltransferase.";  
CC Gene 93:241-247(1990).  
CC [3]  
CC SEQUENCE OF 1-10 FROM N.A., AND CHARACTERIZATION OF THE TWO FORMS.  
CC MEDLINE=92111472; PubMed=1765063;  
CC Ulmanen I., Lundstroem K.;  
CC "Cell-free synthesis of rat and human catechol O-methyltransferase.  
CC Insertion of the membrane-bound form into microsomal membranes in  
CC vitro.";  
CC Eur. J. Biochem. 202:1013-1020(1991).  
CC [4]  
CC X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF SOLUBLE FORM.  
CC MEDLINE=94173341; PubMed=8127373;  
CC Vidgren J., Svensson L.A., Liljas A.;  
CC "Crystal structure of catechol O-methyltransferase.";  
CC Nature 368:354-358(1994).  
CC -!- FUNCTION: Catalyzes the O-methylation, and thereby the  
CC inactivation, of catecholamine neurotransmitters and catechol  
CC hormones. Also shortens the biological half-lives of certain  
CC neuroactive drugs, like L-DOPA, alpha-methyl DOPA and  
CC isoproterenol.  
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + catechol = S-  
CC adenosyl-L-homocysteine + guaiacol.  
CC -!- COFACTOR: Magnesium.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (isoform S-COMT). Type II  
CC membrane protein (isoform MB-COMT).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative initiation;  
CC Comment=2 isoforms, Membrane-bound/MB-COMT (shown here) and  
CC Soluble/S-COMT, are produced by alternative initiation;  
CC -!- PTM: The N-terminus is blocked.  
CC -!- SIMILARITY: TO OTHER MAMMALIAN CATECHOL-O-METHYLTRANSFERASE.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z12651; CAA78276.1; -.

```
DR EMBL: M60754; AAA40892.1; ALT INIT.
DR EMBL: M60753; AAA40881.1; ALT_INIT.
DR PIR: S22090; S22090.
DR FDB; IVID; 11-JUL-96.
DR InterPro; IPR002935; Methyltransf_3.
DR InterPro; IPR000051; SAM_bind.
DR Pfam; PF01596; Methyltransf_3; 1.
KW Transferase; Methyltransferase; Neurotransmitter degradation;
KW Catecholamine metabolism; Transmembrane; Signal-anchor; Magnesium;
KW Alternative initiation; 3D-structure.
FT CHAIN 1 284
FT CATECHOL-O-METHYLTRANSFERASE, ISOFORM
FT MEMBRANE-BOUND.
FT CHAIN 45 264
FT CATECHOL-O-METHYLTRANSFERASE, ISOFORM
FT SOLUBLE.
FT INIT MET 44 44
FT TRANSMEM 3 19
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT ACT_SITE 187 187
FT ACT_SITE 242 242
FT METAL 184 184
FT METAL 212 212
FT METAL 213 213
FT HELIX 48 59
FT TURN 62 63
FT HELIX 65 78
FT TURN 79 79
FT STRAND 104 108
FT TURN 111 112
FT HELIX 114 120
FT TURN 121 122
FT TURN 125 126
FT STRAND 128 133
FT HELIX 136 149
FT TURN 150 150
FT TURN 152 154
FT STRAND 155 159
FT HELIX 162 165
FT HELIX 166 168
FT STRAND 169 173
FT HELIX 178 183
FT STRAND 187 189
FT HELIX 190 199
FT TURN 200 201
FT STRAND 203 211
FT HELIX 214 218
FT TURN 220 228
FT TURN 230 231
FT STRAND 232 240
FT TURN 242 243
FT STRAND 247 255
SQ SEQUENCE 264 AA; 29597 MW; F535DFP49C062854 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 264;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 9 GLLLLALL 16
|||||

RESULT 20
OPDE_PSEAE STANDARD; PRT; 402 AA.
AC Q01602;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription regulatory protein opde.
GN OPDE OR PA2219.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
```

```
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A. / PAO1 / H103;
RC STRAIN=ATCC 15632 / PAO1 / H103;
RA MEDLINE=93051258; PubMed=1427017;
RX Huang H., Siehn R.J., Bellido F., Rawling E., Hancock R.E.W.;
"Analysis of two gene regions involved in the expression of the
RT imipenem-specific, outer membrane porin OprD of Pseudomonas
aeruginosa.";
RL FEWS Microbiol. Lett. 76:267-274(1992).
RN [2]
RP SEQUENCE FROM N.A. / PAO1;
RC STRAIN=ATCC 15632 / PAO1;
RA MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
RA Hickey M.J., Brinkman P.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -!- FUNCTION: REGULATES THE EXPRESSION OF OPD WHICH ENCODES THE
CC IMIPENEM-SPECIFIC PORIN.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO B.SUBTILIS IPA-79D.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: Z14064; CAA78446.1; -.
DR PIR: S23860; S23860.
DR InterPro; IPR007114; MPS.
DR InterPro; IPR005828; Sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00850; MFS; 1.
KW Transcription regulation; Transmembrane; Complete proteome.
FT TRANSMEM 22 42
FT TRANSMEM 60 80
FT TRANSMEM 86 106
FT TRANSMEM 103 128
FT TRANSMEM 147 167
FT TRANSMEM 170 190
FT TRANSMEM 220 240
FT TRANSMEM 256 276
FT TRANSMEM 296 316
FT TRANSMEM 318 338
FT TRANSMEM 348 368
FT TRANSMEM 375 395
SQ SEQUENCE 402 AA; 41592 MW; 0C5701C4AD2FDE16 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 402;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1147 GSTLGLLL 1154
DB 360 GSTLGLLL 367
|||||

RESULT 21
NSMA_HUMAN STANDARD; PRT; 423 AA.
ID NSMA_HUMAN
AC C60906; Q9BWR3;
DT 28-FEB-2003 (Rel. 41, Created)
```



28-FEB-2003 (Rel. 41, Last sequence update)  
10-OCT-2003 (Rel. 42, Last annotation update)  
DE Spingomyelin phosphodiesterase 2 (EC 3.1.4.12) (Neutral  
DE spingomyelinase) (hsMase) (N-SMase) (Lyso-platelet activating factor-  
DE phospholipase C) (Lyso-PAF-PLC).  
GN SMPD2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN [2]  
RN [3]  
SEQUENCE FROM N.A., CHARACTERIZATION, AND TISSUE SPECIFICITY.  
RX MEDLINE=98188255; PubMed=9520418;  
RA Tomiuk S., Hofmann K., Nix M., Zumbansen M., Stoffel W.;  
RT "Cloned mammalian neutral sphingomyelinase: functions in sphingolipid  
RT signaling?";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:3638-3643 (1998).  
RN [2]  
RN [3]  
SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong J.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.X., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smallus D.B.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [3]  
RN [3]  
RX MEDLINE=20076490; PubMed=10608884;  
RA Sawai H., Domae N., Nagai N., Hannun Y.A.;  
RT "Function of the cloned putative neutral sphingomyelinase as  
RT lyso-platelet activating factor-phospholipase C.";  
RL J. Biol. Chem. 274:38131-38139 (1999).  
CC -!- FUNCTION: Converts sphingomyelin to ceramide. Hydrolyze 1-acyl-2-  
CC lyso-sn-glycero-3-phosphocholine (lyso-PC) and 1-O-alkyl-2-lyso-  
CC sn-glycero-3-phosphocholine (lyso-platelet activating factor). The  
CC physiological substrate seems to be Lyso-PAF.  
CC -!- CATALYTIC ACTIVITY: Sphingomyelin + H(2)O = N-acylsphingosine +  
CC choline phosphate.  
CC -!- COFACTOR: Magnesium.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- MISCELLANEOUS: This protein has an optimum pH of 6.5-7.5.  
CC -!- SIMILARITY: Belongs to the neutral sphingomyelinase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AJ222801; CAA10995.1; -;  
CC EMBL; BC000038; AAH00038.1; -;  
CC Genew; HGNC:11121; SMPD2.  
CC MIM; 603498; -;  
CC GO; GO:0005887; C:integral to plasma membrane; TAS.  
CC GO; GO:0004767; P:sphingomyelin phosphodiesterase activity; TAS.

DR GO; GO:0006684; P:sphingomyelin metabolism; TAS.  
DR InterPro; IPR0051135; Exo endo phos.  
DR Pfam; PF03372; Exo endo phos; 1  
KW Hydrolase; Transmembrane; Magnesium  
FT TRANSMEM 330 350 POTENTIAL.  
FT TRANSMEM 354 374 POTENTIAL.  
FT METAL 49 49 MAGNESIUM (BY SIMILARITY).  
FT SITE 180 180 IMPORTANT FOR SUBSTRATE RECOGNITION (BY  
FT SIMILARITY).  
FT ACT SITE 272 272 GENERAL BASE (BY SIMILARITY).  
FT CONFLICT 3 3 L->P (IN REF. 2).  
FT CONFLECT 423 AA; 47592 MW; 05252A923E363171 CRC64;  
SQ  
Query Match 0.7%; Score 8; DB 1; Length 423;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1152 GLILLALL 1159  
Db 334 GLILLALL 341  
RESULT 22  
ENG\_MYCPE  
ID -ENG\_MYCPE STANDARD; PRT; 444 AA.  
AC Q8W5H6; 2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE GTP-binding protein enga.  
DE ENGA OR MYP2290.  
GN Mycoplasma penetrans.  
OS Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OC NCBI\_TaxID=28227;  
RN [1]  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN=HF-2;  
RX MEDLINE=22354719; PubMed=12466555;  
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenzi T., Furuya K.,  
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;  
RT "The complete genomic sequence of Mycoplasma penetrans, an  
RT intracellular bacterial pathogen in humans.";  
RL Nucleic Acids Res. 30:5293-5300 (2002).  
CC -!- FUNCTION: GTPase of unknown physiological role.  
CC -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding  
CC proteins. Enga subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AP004170; BAC44020.1; -;  
CC HAMAP; MF 00195; -; 1.  
DR InterPro; IPR005289; GTP-binding\_dom.  
DR InterPro; IPR002917; MVR\_HSR1.  
DR InterPro; IPR001806; Ras\_trnafrmg.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF01926; MVR\_HSR1; 1.  
DR PRINTS; PR00449; RASTRNSFRMG.  
DR TIGRFAMs; TIGR00650; MG442; 2.  
DR TIGRFAMs; TIGR00231; small GTP; 2.  
KW GTP-binding; Repeat; Complete proteome.  
FT NP\_BIND 8 15 GTP 1 (POTENTIAL).  
FT NP\_BIND 55 59 GTP 1 (POTENTIAL).  
FT NP\_BIND 118 121 GTP 1 (POTENTIAL).  
FT NP\_BIND 179 186 GTP 2 (POTENTIAL).  
FT NP\_BIND 226 230 GTP 2 (POTENTIAL).  
FT NP\_BIND 291 294 GTP 2 (POTENTIAL).  
SQ SEQUENCE 444 AA; 51231 MW; A7B636F43D7CB1BE CRC64;

Query Match 0.7%; Score 8; DB 1; Length 444;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 IIVLDGS 172  
DB 258 IIVLDGS 265

## RESULT 23

HISX\_COREF  
ID HISX\_COREF STANDARD; PRT; 451 AA.  
AC Q8FNZO;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Histidinol dehydrogenase (EC 1.1.1.23) (Hdh)  
GN HISD OR CE2003.  
OS Corynebacterium efficiens.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
OX NCBI\_TaxID=152794;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;  
RX MEDLINE=22723752; PubMed=12840036;  
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,  
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,  
RA Gotojori T.;  
RT "Comparative complete genome sequence analysis of the amino acid  
RT replacements responsible for the thermostability of Corynebacterium  
RL Genome Res. 13:1572-1579 (2003).  
CC -!- FUNCTION: Catalyzes the sequential NAD-dependent oxidations of L-  
CC histidinol to L-histidinolaldehyde and then to L-histidine (by  
CC similarity).  
CC -!- CATALYTIC ACTIVITY: L-histidinol + 2 NAD(+) + H(2)O = L-histidine  
CC + 2 NADH.  
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
CC -!- PATHWAY: Histidine biosynthesis; ninth (last) step.  
CC -!- SIMILARITY: Belongs to the histidinol dehydrogenase family.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AP005220; BAC18813.1; -  
CC HAMAP; MF 01024; -; 1.  
CC InterPro; IPR001692; Histidinol dh.  
CC Pfam; PF00815; Histidinol dh; 1.  
CC PRINTS; PR00083; H0LDHGRNASE  
CC ProDom; PD002880; Histidinol\_dh; 1.  
CC TIGRFAMs; TIGR00069; hisD; 1.  
CC PROSITE; PS00611; HISOL DEHYDROGENASE; FALSE NEG.  
CC Histidine biosynthesis; Oxidoreductase; NAD; Metal-binding; Zinc;  
KW Complete proteome.  
FT ACT\_SITE 332 333 BY SIMILARITY.  
FT ACT\_SITE 333 333 BY SIMILARITY.  
FT METAL 263 263 ZINC (BY SIMILARITY).  
FT METAL 266 266 ZINC (BY SIMILARITY).  
FT METAL 366 366 ZINC (BY SIMILARITY).  
FT METAL 425 425 ZINC (BY SIMILARITY).  
SQ SEQUENCE 451 AA; 47770 MW; 5A2D5AA7B990ECCF CRC64;

Query Match 0.7%; Score 8; DB 1; Length 451;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 DEAAALKDI 341  
DB 402 DEAAALKDI 409

## RESULT 24

D4DR\_HUMAN  
ID D4DR\_HUMAN STANDARD; PRT; 467 AA.  
AC P21917;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE D(4) dopamine receptor (D(2C) dopamine receptor).  
GN DRD4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ALLELE D4.7).  
RC MEDLINE=92310588; PubMed=1319557;  
RA van Tol H.H., Wu C.M., Guan H.C., Ohara K., Bunzow J.R.,  
RA Civelli O., Kennedy J., Seeman P., Niznik H.B., Jovanovic V.;  
RT "Multiple dopamine D4 receptor variants in the human population";  
RL Nature 358:149-152 (1992).  
RN [2]  
RP SEQUENCE FROM N.A. (ALLELE D4.2).  
RC TISSUE=Brain;  
RX MEDLINE=91204054; PubMed=1840645;  
RA van Tol H.H.M., Bunzow J.R., Guan H.-C., Sunahara R.K., Seeman P.,  
RA Niznik H.B., Civelli O.;  
RT "Cloning of the gene for a human dopamine D4 receptor with high  
RT affinity for the antipsychotic clozapine";  
RL Nature 350:610-614 (1991).  
RN [3]  
RP 3D-STRUCTURE MODELING.  
RX MEDLINE=93038586; PubMed=1358063;  
RA Livingstone C.D., Strange P.G., Naylor L.H.;  
RT "Molecular modelling of D2-like dopamine receptors";  
RL Biochem. J. 287:277-282 (1992).  
RN [4]  
RP VARIANT GLY-194.  
RX MEDLINE=95243275; PubMed=7726213;  
RA Seeman P., Ulpian C., Chouinard G., van Tol H.H.M., Dwosh H.,  
RA Lieberman P.A., Sminovitch K., Liu I.S.C., Waye J., Voruganti P.,  
RA Rudson C., Serjeant G.R., Masibay A.S., Seeman M.V.;  
RT "Dopamine D4 receptor variant, D4-glycine-194, in Africans, but not in  
RT Caucasians: no association with schizophrenia";  
RL Am. J. Med. Genet. 54:384-390 (1994).  
CC -!- FUNCTION: This is one of the five types (D1 to D5) of receptors  
CC for dopamine. The activity of this receptor is mediated by G  
CC proteins which inhibit adenylyl cyclase.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- POLYMORPHISM: The number of repeats is highly polymorphic and  
CC varies among different alleles. The sequence shown is that of  
CC allele D4.7.  
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; L12398; AAB59386.1; -  
CC PIR; S15079; DYH4.  
CC Genew; HGNC:3025; DRD4.  
CC MIM; 128452; -  
CC GO; GO:0005887; C:integral to plasma membrane; TAS.  
CC GO; GO:0004952; F:dopamine receptor activity; TAS.

DR GO; GO:0007166; P:cell surface receptor linked signal transdu. ...; TAS.  
DR GO; GO:0007212; P:dopamine receptor signaling pathway; TAS.  
DR GO; GO:0007268; P:synaptic transmission; TAS.  
DR InterPro; IPR000276; GPCR\_Rhodopsin.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PS00237; GPCR\_Rhodopsin.  
DR PROSITE; PS00237; G-PROTEIN RECEPTOR F1 1; 1.  
DR PROSITE; PS00262; G-PROTEIN RECEPTOR F1 2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW Multigene family; Polymorphism; Repeat.  
FT DOMAIN 1 37 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 38 60 1 (POTENTIAL).  
FT DOMAIN 61 70 2 (POTENTIAL).  
FT TRANSMEM 71 93 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 94 109 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 110 131 3 (POTENTIAL).  
FT DOMAIN 132 151 4 (POTENTIAL).  
FT TRANSMEM 152 175 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 176 191 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 192 213 5 (POTENTIAL).  
FT DOMAIN 214 394 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 395 417 6 (POTENTIAL).  
FT DOMAIN 418 426 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 427 449 7 (POTENTIAL).  
FT DOMAIN 450 467 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 468 486 7 X 16 AA APPROXIMATE TANDEM REPEATS.  
FT REPEAT 249 360 1.  
FT REPEAT 265 280 2.  
FT REPEAT 281 296 3.  
FT REPEAT 297 312 4.  
FT REPEAT 313 328 5.  
FT REPEAT 329 344 6.  
FT REPEAT 345 360 7.  
FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT DISULFID 108 195 BY SIMILARITY.  
FT VARIANT 194 194 V -> G (in Afro-Caribbeans;  
dbSNP:1800443).  
FT VARIANT 265 344 Missing (in allele D4.2).  
FT VARIANT 281 328 Missing (in allele D4.4).  
FT VARIANT 329 329 Missing (in allele D4.4).  
FT VARIANT 332 332 Missing (in allele D4.4).  
FT VARIANT 332 332 G -> S (in allele D4.4).  
FT SEQUENCE 467 AA; 48360 MW; B6PF2E09269A02AF CRC64;  
Query Match 0.7%; Score 8; DB 1; Length 467;  
Best Local Similarity 100.0%; Pred. No. 21; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1154 LLLALLVL 1161  
Db 81 LLLALLVL 88  
RESULT 25  
SIL7\_HUMAN STANDARD; PRT; 467 AA.  
AC Q9Y286; Q9Y286; Q9Y286; Q9Y286; Q9Y286;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Sialic acid binding Ig-like lectin 7 precursor (Siglec-7) (Q9Y286)  
DE membrane protein (Adhesion inhibitory receptor molecule-1) (AIRM-1)  
DE (p75) (D-siglec).  
GN SIGLEC7 OR AIRM1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Dendritic cell;  
RX MEDLINE=20036547; PubMed=10567377;  
RA Nicoll G., Ni J., Liu D., Klennerman P., Munday J., Dubock S.,  
RA Mattei M.-G., Crocker P.R.;  
RT "Identification and characterization of a novel siglec, siglec-7,  
RT expressed by human natural killer cells and monocytes.";  
J. Biol. Chem. 274:34089-34095(1999).  
[2]  
RN SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), PHOSPHORYLATION, AND  
RP INTERACTION WITH PTPN6.  
RC TISSUE=Lymphoid;  
RX MEDLINE=99429884; PubMed=10499918;  
RA Falco M., Biassoni R., Bottino C., Vitale M., Sivori S.,  
RA Augugliaro R., Moretta L., Moretta A.;  
RT "Identification and molecular cloning of p75/AIRM1, a novel member of  
RT the sialoadhesin family that functions as an inhibitory receptor in  
RT human natural killer cells.";  
J. Exp. Med. 190:793-802(1999).  
[3]  
RN SEQUENCE FROM N.A. (ISOFORM 2).  
RP TISSUE=Dendritic cell;  
RA Zhang W., Wan T., Cao X.;  
RT "Characterization of a novel siglec from dendritic cells.";  
Submitted (AUG-1999) to the EMBL/GenBank/DBSJ databases.  
[4]  
RN SEQUENCE FROM N.A. (ISOFORM 2).  
RP MEDLINE=20230158; PubMed=10764831;  
RA Angata T., Varki A.;  
RT "Siglec-7, a sialic acid-binding lectin of the immunoglobulin  
RT superfamily.";  
RG Glycobiology 10:431-438(2000).  
[5]  
RN FUNCTION.  
RP MEDLINE=20079612; PubMed=10611343;  
RA Vitale C., Romagnani C., Falco M., Ponte M., Vitale M., Moretta A.,  
RA Baggiolupo A., Moretta L., Mingari M.C.;  
RT "Engagement of p75/AIRM1 or CD33 inhibits the proliferation of normal  
RT or leukemic myeloid cells.";  
Proc. Natl. Acad. Sci. U.S.A. 96:15091-15096(1999).  
[6]  
RN DISIALOGLANGLIOSIDE BINDING.  
RX MEDLINE=21286126; PubMed=11389909;  
RA Ito A., Handa K., Withers D.A., Satoh M., Hakomori S.;  
RT "Binding specificity of siglec7 to disialogangliosides of renal cell  
RT carcinoma: possible role of disialogangliosides in tumor  
RT progression.";  
FEBS Lett. 498:116-120(2001).  
CC -!- FUNCTION: Putative adhesion molecule that mediates sialic-acid  
CC dependent binding to cells. Preferentially binds to alpha2,3 and  
CC alpha2,6-linked sialic acid. Also binds disialogangliosides  
CC (disialogangliosyl globoside, disialyl lactotetraosylceramide and  
CC disialyl GalNAc lactotetraosylceramide). The sialic acid  
CC recognition site may be masked by cis interactions with sialic  
CC acids on the same cell surface. In the immune response, may act as  
CC an inhibitory receptor upon ligand induced tyrosine  
CC phosphorylation by recruiting cytoplasmic phosphatase(s) via their  
CC SH2 domain(s) that block signal transduction through  
CC dephosphorylation of signaling molecules. Mediates inhibition of  
CC natural killer cells cytotoxicity. May play a role in hemopoiesis.  
CC Inhibits differentiation of CD34+ cell precursors towards  
CC myelomonocytic cell lineage and proliferation of leukemic myeloid  
CC cells (in vitro).  
CC -!- SURUNIT: Interacts with PTPN6/SHP-1 upon phosphorylation.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=4;  
CC Name=1; Synonyms=AIMM-1b;  
CC IsoId=Q9Y286-1; Sequence=Displayed;  
CC Name=2; Synonyms=AIMM-2;  
CC IsoId=Q9Y286-2; Sequence=VSP\_002555;  
CC Name=3; Synonyms=AIMM-3;  
CC IsoId=Q9Y286-3; Sequence=VSP\_002556, VSP\_002558;  
CC

Note=No experimental confirmation available;  
 Name=4;  
 IsoId=Q9Y296-4; Sequence=VSP\_002557, VSP\_002558;  
 Note=No experimental confirmation available;  
 TISSUE SPECIFICITY: Predominantly expressed by resting and activated natural killer cells and at lower levels by granulocytes and monocytes. High expression found in placenta, liver, lung, spleen, and peripheral blood leukocytes.  
 DOMAIN: Contains 1 copy of a cytoplasmic motif that is referred to as the immunoreceptor tyrosine-based inhibitor motif (ITIM). This motif is involved in downmodulation of cellular responses. The phosphorylated ITIM motif binds to the SH2 domain of FPN6/SHP-1.  
 PTM: Tyrosine phosphorylated.  
 SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLESC (sialic acid binding Ig-like lectin) family.  
 SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.  
 CAUTION: Ref.3 sequence differs from that shown due to a frameshift in position 406.  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 EMBL; AF170485; AAF12759.1; -  
 EMBL; AJ007395; CAB46011.1; -  
 EMBL; AJ130710; CAB51126.1; -  
 EMBL; AJ130711; CAB51127.1; -  
 EMBL; AJ130712; CAB51128.1; -  
 EMBL; AJ130713; CAB51129.1; -  
 EMBL; AF178981; AAF43446.1; ALT\_FRAME.  
 EMBL; AF193441; AAF06790.1; -  
 Genbank; HGNC:10876; SIGLEC7.  
 MIM: 604410; -  
 GO; GO:0005887; C-integral to plasma membrane; TAS.  
 GO; GO:0005530; P-lectin; TAS.  
 GO; GO:0004872; F-receptor activity; TAS.  
 InterPro; IPR007110; Ig-like.  
 Pfam; PF00047; Ig; 3.  
 SMART; SM00408; IGC2; 1.  
 PROSITE; PS00335; IG\_LIKE; 2.  
 Cell adhesion; lectin; Transmembrane; Signal; Glycoprotein;  
 Immunoglobulin domain; Repeat; Phosphorylation; Alternative splicing.  
 SIGNAL 1 18  
 CHAIN 19 467  
 DOMAIN 19 353  
 TRANSMEM 354 376  
 DOMAIN 377 467  
 DOMAIN 39 122  
 DOMAIN 150 233  
 DOMAIN 240 336  
 SITE 435 440  
 DISULFID 46 106  
 DISULFID 168 217  
 DISULFID 276 320  
 CARBOHYD 105 105  
 CARBOHYD 142 142  
 CARBOHYD 165 165  
 CARBOHYD 229 229  
 CARBOHYD 235 235  
 CARBOHYD 242 242  
 CARBOHYD 260 260  
 CARBOHYD 334 334  
 VARSPLIC 145  
 ALTHENILPGLTSCGFLNLTCSVPACOSTPWNW  
 GTVSPLHPSTTSSVLTLPQPHGHTSLTCQVLPAGV  
 TTRRTQLNVSYS -> D (in isoform 2).  
 /FTid=VSP\_002555.

FT VARSPLIC 145 145 A->B (in isoform 3).  
 FT VARSPLIC 145 145 /FTid=VSP\_002556.  
 FT VARSPLIC 145 145 A->G (in isoform 4).  
 FT VARSPLIC 146 467 /FTid=VSP\_002557.  
 FT VARSPLIC 146 467 Missing (in isoform 3 and isoform 4).  
 FT CONFLICT 42 42 V->A (IN REF. 4).  
 FT CONFLICT 42 42 V->A (IN REF. 4).  
 SQ SEQUENCE 467 AA; 51142 MW; 8AFB4462B001F52 CRC64;  
 Query Match 0.7%; Score 8; DB 1; Length 467;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 379 GVLLGAVG 386  
 DB 350 GVLLGAVG 357  
 RESULT 26  
 ID\_CPBK MOUSE STANDARD; PRT; 470 AA.  
 AC Q62397;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome P450 2B20 (EC 1.14.14.1) (CYP2B20) (P24) (Fragment).  
 GN CYP2B20.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Liver;  
 EX MEDLINE=96428606; PubMed=8831708;  
 RA Daron M., Pautrel A., Marc N., Guillouzo A., Corcos L.;  
 RT Isolation of a new mouse cDNA clone: hybrid form of Cytochrome P450  
 2B10 and NADPH-cytochrome P450 oxidoreductase.;  
 RL Biochem. Biophys. Res. Commun. 226:900-905(1996).  
 CC !- FUNCTION: Cytochromes P450 are a group of heme-thiolate  
 CC monooxygenases. In liver microsomes, this enzyme is involved in an  
 CC NADPH-dependent electron transport pathway. It oxidizes a variety  
 CC of structurally unrelated compounds, including steroids, fatty  
 CC acids, and xenobiotics.  
 CC !- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
 CC oxidized flavoprotein + H(2)O.  
 CC !- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC !- TISSUE SPECIFICITY: EXPRESSED IN THE LIVER AS WELL AS IN KIDNEY,  
 CC LUNG AND INTESTINE.  
 CC !- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; X99715; CRA68051.1; ALT\_INIT.  
 DR HSP; P00179; 1D76.  
 DR MGD; MGI:1202389; Cyp2b20.  
 DR InterPro; IPR001128; Cytochrome P450.  
 DR InterPro; IPR008068; EP450\_CYP2B.  
 DR Pfam; PF00067; P450; 1  
 DR PRINTS; PR01685; EP450ICYP2B.  
 DR PROSITE; PS00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum; Phosphorylation.  
 FT MOD\_RES 128 128 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
 FT METAL 436 436 IRON (HEME AXIAL LIGAND).  
 FT NON\_TER 470 470

```
SQ SEQUENCE 470 AA; 53357 MW; 8B9CF3E2EA622642 CRC64;
Query Match 0.7%; Score 8; DB 1; Length 470;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLLV 1160
DB 6 LLLLLLLV 13

RESULT 27
XVLC_PSEPU STANDARD; PRT; 487 AA.
ID XVLC_PSEPU STANDARD; PRT; 487 AA.
AC P43503;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Benzaldehyde dehydrogenase [NAD+] (EC 1.2.1.28).
GN XVLC
OS Pseudomonas putida.
OG Plasmid TOL pMW0, and plasmid TOL pMW3.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.
RC PLASMID-TOL pMW0; PubMed=7868591;
RX MEDLINE=95173094;
RA Inoue J., Shaw J.P., Reik M., Harayama S.;
RT "Overlapping substrate specificities of benzaldehyde dehydrogenase
(the xylC gene product) and 2-hydroxymuconic semialdehyde
dehydrogenase (the xylG gene product) encoded by TOL plasmid pMW0 of
Pseudomonas putida";
RT Pseudomonas putida";
RL J. Bacteriol. 177:1196-1201(1995).
RN [2]
RP SEQUENCE OF 1-53.
RC PLASMID-TOL pMW53;
RX MEDLINE=91113163; PubMed=1989592;
RA Chalmers R.M., Keen J.N., Fawson C.A.;
RT "Comparison of benzyl alcohol dehydrogenases and benzaldehyde
dehydrogenases from the benzyl alcohol and mandelate pathways in
Acinetobacter calcoaceticus and from the TOL-plasmid-encoded toluene
pathway in Pseudomonas putida. N-terminal amino acid sequences, amino
acid compositions and immunological cross-reactions.";
RL Biochem. J. 273:99-107(1991).
CC -1- CATALYTIC ACTIVITY: Benzaldehyde + NAD(+) + H(2)O = benzoate +
NADH.
CC -1- SUBUNIT: Homotetramer.
CC -1- MISCELLANEOUS: Optimal pH is 9.0.
CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: U15151; AAA66218.1;
CC EMBL: D63341; BAA09661.1;
CC PIR: T47107; T47107.
CC RSP; P51977; 1bXS.
CC InterPro: IPR002086; Aldehyde_dehydr.
CC Pfam: PF00171; aldedh; 1.
CC PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; 1.
CC PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD; Plasmid.
FT NP_BIND 232 237 NAD (ADP PART) (BY SIMILARITY).
FT ACT_SITE 254 254 BY SIMILARITY.
FT ACT_SITE 288 288 BY SIMILARITY.
SQ SEQUENCE 487 AA; 51897 MW; 093CB3E9487AF384 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 626 ALGNVAIL 633
DB 171 ALGNVAIL 178

RESULT 28
CPBI_RAT
ID CPBI_RAT STANDARD; PRT; 491 AA.
AC P00176;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2B1 (EC 1.14.14.1) (CYP1B1) (P450-B) (P450-PB1 and
P450-PB2) (P450-LM2).
GN CYP2B1 OR CYP2B-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE OF 6-491 FROM N.A. (ISOZYME PB1 AND PB2).
RX MEDLINE=8222224; PubMed=6953431;
RA Fujii-Kuriyama Y., Mizukami Y., Kawajiri K., Sogawa K., Muramatsu M.;
RT "Primary structure of a cytochrome P-450: coding nucleotide sequence
of phenobarbital-inducible cytochrome P-450 cDNA from rat liver.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2793-2797(1982).
RN [2]
RP REVISIONS TO 166; 292 AND 378 (ISOZYME PB1 AND PB2).
RA Fujii-Kuriyama Y., Mizukami Y., Kawajiri K., Sogawa K., Muramatsu M.;
RL Proc. Natl. Acad. Sci. U.S.A. 79:5443-5443(1982).
RN [3]
RP SEQUENCE OF 1-22.
RX MEDLINE=7919411; PubMed=109438;
RA Botelho L.H., Ryan D.E., Levin W.;
RT "Amino acid compositions and partial amino acid sequences of three
highly purified forms of liver microsomal cytochrome P-450 from rats
treated with polychlorinated biphenyls, phenobarbital, or 3-
methylcholanthrene.";
RL J. Biol. Chem. 254:5635-5640(1979).
RN [4]
RP PHOSPHORYLATION.
RX MEDLINE=90059885; PubMed=2583091;
RA Fyverin W., Taniguchi H.;
RT "Phosphorylation of hepatic phenobarbital-inducible cytochrome
P-450";
RL EMBL J. 8:3003-3010(1989).
CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate
monooxygenases. In liver microsomes, this enzyme is involved in an
NADPH-dependent electron transport pathway. It oxidizes a variety
of structurally unrelated compounds, including steroids, fatty
acids, and xenobiotics.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- INDUCTION: By phenobarbital.
CC -1- PTM: Phosphorylation is accompanied by a decrease in enzyme
activity.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: J00719; AAA41024.1;
CC DR
```

```

DR EMBL; M37134; AAC42028.1; -.
DR PIR; A00176; O4RTTB.
DR PIR; A21162; O4RTP2.
DR HSSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01685; EP450ICYP2B.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Phosphorylation.
FT MOD_RES 128 128 PHOSPHORYLATION (BY PKA).
FT METAL 436 436 IRON (HEME AXIAL LIGAND).
FT VARIANT 303 303 S -> G (IN ISOZYME PB2).
FT VARIANT 321 322 AE -> IV (IN ISOZYME PB2).
FT VARIANT 337 337 L -> P (IN ISOZYME PB2).
FT VARIANT 339 339 T -> S (IN ISOZYME PB2).
FT VARIANT 344 344 S -> T (IN ISOZYME PB2).
SQ SEQUENCE 491 AA; 55933 MW; 74615501AD5497DD CRC64;

Query Match 0.7%; Score 8; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
Db 6 LLLALLLV 13

RESULT 29
CPB2_RAT STANDARD; PRT; 491 AA.
AC P04187; Q64582;
DT 20-MAR-1987 (Rel. 04, Created)
DT 15-JUL-1998 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2B2 (SC 1.14.14.1) (CYP11B2) (P450B) (P450 PB4).
GN CYP2B2 OR CYP2B-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83247397; PubMed=6306654;
RA Mizukami Y., Sogawa K., Suwa Y., Muramatsu M., Fujii-Kuriyama Y.;
RT "Gene structure of a phenobarbital-inducible cytochrome P-450 in rat liver."
RL Proc. Natl. Acad. Sci. U.S.A. 80:3958-3962(1983).
RN [2]
RP SEQUENCE
RX MEDLINE=86059379; PubMed=3877725;
RA Frey A.B., Waxman D.J., Kreibich G.;
RT "The structure of phenobarbital-inducible rat liver cytochrome P-450 isoenzyme PB-4. Production and characterization of site-specific antibodies."
RL J. Biol. Chem. 260:15253-15265(1985).
RN [3]
RP SEQUENCE OF 168-491 FROM N.A.
RX MEDLINE=84159487; PubMed=6689485;
RA Phillips I.R., Shephard E.A., Ashworth A., Rabin B.R.;
RT "Cloning and sequence analysis of a rat liver cDNA coding for a phenobarbital-inducible microheterogenous cytochrome P-450 variant: regulation of its messenger level by xenobiotics."
RL Gene 26:41-52(1983).
RN [4]
RP SEQUENCE OF 281-491 FROM N.A.
RX MEDLINE=83291091; PubMed=6688421;
RA Kumar A., Raphael C., Adesnik M.;
RT "Cloned cytochrome P-450 cDNA. Nucleotide sequence and homology to multiple phenobarbital-induced mRNA species."
RL J. Biol. Chem. 258:111280-11284(1983).

```

```

RN [5]
RP ERRATUM.
RA Kumar A., Raphael C., Adesnik M.;
RL J. Biol. Chem. 259:6039-6039(1984).
RN [6]
RP MEDLINE=86205943; PubMed=3458196;
RA Atchison M.L., Adesnik M.;
RT "Gene conversion in a cytochrome P-450 gene family."
RL Proc. Natl. Acad. Sci. U.S.A. 83:2300-2304(1986).
RN [7]
RP SEQUENCE OF 385-491 FROM N.A.
RX MEDLINE=84153837; PubMed=6322758;
RA Affolter M., Anderson A.;
RT "Segmental homologies in the coding and 3' non-coding sequences of rat liver cytochrome P-450b and P-450b cDNAs and cytochrome P-450e-like genes."
RL Biochem. Biophys. Res. Commun. 118:655-662(1984).
RN [8]
RP SEQUENCE OF 1-20 FROM N.A.
RX MEDLINE=88273074; PubMed=2839467;
RA Hashimoto T., Matsumoto T., Nishizawa M., Kawabata S.,
RT "A mutant rat strain deficient in induction of a phenobarbital-inducible form of cytochrome P-450 in liver microsomes."
RL J. Biochem. 103:487-492(1988).
RN [9]
RP PHOSPHORYLATION.
RX MEDLINE=90059885; PubMed=2583091;
RA Pyerin W., Taniguchi H.;
RT "Phosphorylation of hepatic phenobarbital-inducible cytochrome P-450."
RL EMBO J. 8:3003-3010(1989).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate monooxygenases. In liver microsomes, this enzyme is involved in an NADPH-dependent electron transport pathway. It oxidizes a variety of structurally unrelated compounds, including steroids, fatty acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH + oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: By phenobarbital.
CC -!- PTM: Phosphorylation is accompanied by a decrease in enzyme activity.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR EMBL; J00728; AAA41056.1; -.
DR EMBL; J00720; AAA41056.1; JOINED.
DR EMBL; J00721; AAA41056.1; JOINED.
DR EMBL; J00722; AAA41056.1; JOINED.
DR EMBL; J00723; AAA41056.1; JOINED.
DR EMBL; J00724; AAA41056.1; JOINED.
DR EMBL; J00725; AAA41056.1; JOINED.
DR EMBL; J00726; AAA41056.1; JOINED.
DR EMBL; K00996; AAA41029.1; -.
DR EMBL; K01626; AAA41037.1; -.
DR EMBL; K01721; AAA41026.1; -.
DR EMBL; D00250; BAA00181.1; -.
DR EMBL; M13234; AAA41057.1; -.
DR PIR; A21162; O4RTP2.
DR HSSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008068; EP450_CYP2B.

```

```

DR PRINTS; PR01685; EP450ICYP2B..
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; FALSE NEG.
DR OXIDOREDUCTASE; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Phosphorylation.
FT MOD_RES 128 128 PCSPHORylation (BY PKA) (BY SIMILARITY).
FT FT 445 445 IRON (HEME AXIAL LIGAND).
FT METAL 445 445 IRON (HEME AXIAL LIGAND).
SQ SEQUENCE 500 AA; 56743 MW; F560A00D80DFBA94 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLLV 1160
Db 6 LLLLLLLV 13

RESULT 31
VEI_BP2V ID VEI_BP2V STANDARD; PRT; 604 AA.
AC F11298;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE DE Replication protein E1.
GN E1.
OS Bovine papillomavirus type 2.
OC Viruses; GSDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10560;
RN [1]
RP SEQUENCE FROM N.A.
RA Groff D.E., Mitra R., Lancaster W.D.;
RL Submitted (MAY-1988) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF
CC VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
CC PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
CC CONTAINS BINDING SITES FOR BOTH PROTEINS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M20219; AAA66833.1; -.
CC PIR; C31169; W1WLE2.
CC InterPro; IPR001177; Papillom_E1.
CC DR PFam; PF00519; E1; 1.
CC DR Pfam; PF00524; E1_N; 1.
CC KW Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
CC Nuclear protein.
CC NP_BIND 432 439 ATP (POTENTIAL).
CC SEQUENCE 604 AA; 68077 MW; D2D7036AD8E8A9DD CRC64;

Query Match 0.7%; Score 8; DB 1; Length 604;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 927 LAAGSDSN 934
Db 336 LAAGSDSN 343

RESULT 32
VEI_SPV1 ID VEI_SPV1 STANDARD; PRT; 605 AA.
AC F03116; Q9WWH1;
DT 21-JUL-1986 (Rel. 01, Created)

```

```

DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Replication protein E1.
GN E1.
OS Bovine papillomavirus type 1.
OC Viruses; deDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10559;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=83012974; PubMed=6289124;
RA Chen E.Y., Howley P.M., Levinson A.D., Seeburg P.H.;
RT "The primary structure and genetic organization of the bovine
RL papillomavirus type 1 genome.";
RN Nature 299:529-534(1982).
RN [2]
RN REQUIREMENT FOR REPLICATION.
RX MEDLINE=91122053; PubMed=1846806;
RA Ustav M., Stenlung A.;
RT "Transient replication of BPV-1 requires two viral polypeptides
RL encoded by the E1 and E2 open reading frames.";
RN EMBO J. 10:449-457(1991).
RN [3]
RN CHARACTERIZATION.
RX MEDLINE=93281701; PubMed=8389467;
RA Yang L., Mohr I., Fouts E., Lim D.A., Nohale M., Botchan M.;
RT "The E1 protein of bovine papilloma virus 1 is an ATP-dependent DNA
RL helicase";
RN Proc. Natl. Acad. Sci. U.S.A. 90:5086-5090(1993).
CC -1- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF
CC VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
CC PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
CC CONTAINS BINDING SITES FOR BOTH PROTEINS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X02346; CAB46511.1; -
CC PIR; A03663; WIMLEB.
CC PDB; 1F08; 16-MAY-01.
CC TRANSFAC; T00203; -
CC InterPro; IPR001177; Papillom_E1.
CC Pfam; PF00519; E1_1.
CC Pfam; PF00524; E1_N; 1.
CC Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
KW Nuclear protein; 3D-structure.
FT NP BIND 433 440 ATP (POTENTIAL).
SQ SEQUENCE 605 AA; 68190 MW; C8400B7B8F60620B CRC64;

Query Match 0.7%; Score 8; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 927 LAAGSDSN 934
Db 337 LAAGSDSN 344
|||||
|||||

RESULT 33
ID TWS6_HUMAN STANDARD; PRT; 811 AA.
AC Q8IU80; Q8IU82; Q8IXV8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transmembrane protease, serine 6 (EC 3.4.21.-) (Matrptase-2).
GN TMPRSS6.

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RA Hooper J.D., Quigley J.P.;
RT "TMPRSS6, a new type II transmembrane serine protease.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208;
RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smit J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson G.,
RA Evans P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA Gilbert J.G.R., Goward M.B., Graham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcott R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,
RA McElay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Shira H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shitani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyma S.,
RA Roe B.A., Chen P., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang P., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA Phan Q., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
RA Peyrard M., Kedra D., Seroussi E., Fransoni I., Tapia I., Bruder C.E.,
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA Khan A.S., Lane L., Tilahun Y., Wright H.;
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
RN [3]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Souffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

```





```

CC CC ----- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL; U12683; AAA99668.1; --
CC DR EMBL; L10191; AAA16249.1; --
CC DR PIR; I51527; I51527.
CC DR HSPP; P06756; IJUV2.
CC DR InterPro; IPR000413; Integrin_alpha.
CC DR Pfam; PF01839; FG-GAP; 4.
CC DR Pfam; PF00357; Integrin_A; 1.
CC DR SMART; SMC0191; Int_alpha; 5.
CC DR PROSITE; PS00242; INTEGRIN ALPHA; 1.
CC DR Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC KW Signal; Repeat.
CC PT SIGNAL; Repeat.
CC FT CHAIN 33 1050
CC FT CHAIN 33 932
CC FT CHAIN 933 1050
CC FT DOMAIN 33 996
CC FT TRANSMEM 997 1022
CC FT DOMAIN 1023 1050
CC FT REPEAT 48 110
CC FT REPEAT 120 188
CC FT REPEAT 189 246
CC FT REPEAT 259 312
CC FT REPEAT 313 373
CC FT REPEAT 379 438
CC FT REPEAT 442 494
CC FT CA_BIND 324 332
CC FT CA_BIND 390 398
CC FT CA_BIND 454 462
CC FT SITE 1025 1029
CC FT DISULFID 90 99
CC FT DISULFID 145 166
CC FT DISULFID 182 195
CC FT DISULFID 502 513
CC FT DISULFID 519 575
CC FT DISULFID 636 642
CC FT DISULFID 708 721
CC FT DISULFID 862 910
CC FT DISULFID 917 922
CC FT CARBOHYD 75 75
CC FT CARBOHYD 95 95
CC FT CARBOHYD 98 98
CC FT CARBOHYD 172 172
CC FT CARBOHYD 287 287
CC FT CARBOHYD 297 297
CC FT CARBOHYD 306 306
CC FT CARBOHYD 507 507
CC FT CARBOHYD 515 515
CC FT CARBOHYD 521 521
CC FT CARBOHYD 600 600
CC FT CARBOHYD 649 649
CC FT CARBOHYD 714 714
CC FT CARBOHYD 763 763
CC FT CARBOHYD 861 861
CC SQ SEQUENCE 1050 AA; 115961 MW; 10ED9615358BD918 CRC64;
CC
CC Query Match 0.7%; Score 8; DB 1; Length 1050;
CC Best Local Similarity 100.0%; Pred. No. 41;
CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 541 RFGSSIAS 548
CC DB 380 RFGSSIAS 387

```

## RESULT 35

```

ITAL DROME STANDARD; PRT; 1146 AA.
AC Q2447; Q9YVF6;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-PS1 precursor (Position-specific antigen 1, alpha
DE chain) (Protein multiple edematous wings).
GN MEW OR CG1771.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Oregon-R;
RX MEDLINE=94059764; PubMed=8240969;
RA Wehrli M., DiAntonio A., Fearnley I.M., Smith R.J., Wilcox M.;
RT "Cloning and characterization of alpha PS1, a novel Drosophila
RT melanogaster integrin.";
RL Mech. Dev. 43:21-36(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pictman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Stden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
CC Science 287:2185-2195(2000).
CC -!- FUNCTION: INTEGRIN ALPHA-PS1/BETA-PS IS A RECEPTOR FOR LAMININ.
CC -!- SUBUNIT: HETEROIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC -!- SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A
CC DISULFIDE BOND. ALPHA-PS1 ASSOCIATES WITH BETA-PS.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous

```



RT CR3 (CD11b/CD18).";  
 RL Cell 80:631-638(1995).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.  
 RX MEDLINE=96363671; PubMed=8747460;  
 RA Lee J.O., Bankston L.A., Arnaut M.A., Liddington R.C.;  
 RT "Two conformations of the integrin A-domain (I-domain): a pathway for  
 activation?";  
 RL Structure 3:1333-1340(1995).  
 RN [11]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.  
 RX MEDLINE=98362595; PubMed=9687375;  
 RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,  
 Fairbanks M.B., Finkel B.C., Garlick R.L., Heinrichson R.L.,  
 Horton N.C., Kelley L.L., Midner A.M., Moon J.B., Mott J.E.,  
 Mutschler V.I., Tomich C.S., Watenpaugh K.D., Wiley V.H.;  
 RT "Cation binding to the integrin CD11b I domain and activation model  
 assessment.";  
 RL Structure 6:923-935(1998).  
 RN [12]  
 RP 3D-STRUCTURE MODELING OF 17-616.  
 RX MEDLINE=98226734; PubMed=9560195;  
 RA Oxvig C., Springer T.A.;  
 RT "Experimental support for a beta-propeller domain in integrin alpha-  
 subunits and a calcium binding site on its lower surface.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875(1998).  
 CC -!- FUNCTION: INTEGRIN ALPHA-M/BETA-2 IS IMPLICATED IN VARIOUS  
 ADHESIVE INTERACTIONS OF MONOCYTES, MACROPHAGES AND GRANULOCYTES  
 AS WELL AS IN MEDIATING THE UPTAKE OF COMPLEMENT-COATED PARTICLES.  
 IT IS IDENTICAL WITH CR-3, THE RECEPTOR FOR THE IC3B FRAGMENT OF  
 THE THIRD COMPLEMENT COMPONENT. IT PROBABLY RECOGNIZES THE R-G-D  
 PEPTIDE IN C3B. INTEGRIN ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR  
 FIBRINOGEN, FACTOR X AND ICAM1. IT RECOGNIZES P1 AND P2 PEPTIDES  
 OF FIBRINOGEN GAMMA CHAIN.  
 CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-M  
 ASSOCIATES WITH BETA-2.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND  
 GRANULOCYTES.  
 CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS  
 WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- DATABASE: NAME=PROV; NOTE=CD guide CDlib entry;  
 WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cdlib.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: S52180; AAB24821.1; JOINED.  
 DR EMBL: S52181; AAB24821.1; JOINED.  
 DR EMBL: S52184; AAB24821.1; JOINED.  
 DR EMBL: S52189; AAB24821.1; JOINED.  
 DR EMBL: S52191; AAB24821.1; JOINED.  
 DR EMBL: S52192; AAB24821.1; JOINED.  
 DR EMBL: S52203; AAB24821.1; JOINED.  
 DR EMBL: S52212; AAB24821.1; JOINED.  
 DR EMBL: S52213; AAB24821.1; JOINED.  
 DR EMBL: S52216; AAB24821.1; JOINED.  
 DR EMBL: S52219; AAB24821.1; JOINED.  
 DR EMBL: S52220; AAB24821.1; JOINED.  
 DR EMBL: S52221; AAB24821.1; JOINED.  
 DR EMBL: S52222; AAB24821.1; JOINED.  
 DR EMBL: S52226; AAB24821.1; JOINED.  
 DR EMBL: M76724; AAB24821.1; JOINED.  
 DR EMBL: M84477; AAB24821.1; JOINED.  
 DR PIR: A31108; RWHU1B.  
 DR PDB: 1A8X; 17-JUN-98.  
 DR PDB: 1BHO; 18-NOV-98.  
 DR PDB: 1BHO; 18-NOV-98.  
 DR PDB: 1IDN; 25-NOV-98.  
 DR PDB: 1IDO; 01-AUG-96.  
 DR PDB: 1JLW; 11-JAN-97.  
 DR PDB: 1MIU; 07-AUG-02.  
 DR Genew; HGNC:6149; ITGAM.  
 DR MIM; 120980; -.  
 DR GO; GO:0008305; C: integrin complex; TAS.  
 DR GO; GO:0004895; P: cell adhesion receptor activity; TAS.  
 DR GO; GO:0007155; P: cell adhesion; TAS.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWF\_A.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00357; Integrin\_A; 1.  
 DR Pfam; PF00092; vwa; 1.  
 DR PRINTS; PR01185; INTEGRINA.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00191; Int\_alpha; 4.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS00234; VWFA; 1.  
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;  
 KW Signal; 3D-structure; Repeat; Magnesium; Calcium.  
 FT SIGNAL 1 16  
 FT CHAIN 17 1152 INTEGRIN ALPHA-M.  
 FT DOMAIN 17 1104 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1105 1128 POTENTIAL.  
 FT DOMAIN 1129 1152 CYTOPLASMIC (POTENTIAL).  
 Query Match 0.7%; Score 8; DB 1; Length 1152;  
 Best Local Similarity 100.0%; Pred. No. 44;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1151 GGLLLAL 1158  
 DB 1115 GGLLLAL 1122  
 RESULT 37  
 DD37\_HUMAN STANDARD; PRT; 1157 AA.  
 ID DD37\_HUMAN  
 AC Q81Y37; Q9BU17; Q9P211;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Probable ATP-dependent helicase DHX37 (DEAH-box protein 37).  
 GN DHX37 OR DDH37 OR KIAA1517.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Brain, and Pubmednum;  
RX MEDLINE=2238257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny N.J., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fabsy J., Helton B., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE OF 178-1157 FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=20277482; PubMed=10819331;  
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XVII.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 7:143-150(2000).  
CC -!- SIMILARITY: Belongs to the DEAD box helicase family. DEAD  
CC subfamily.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [licenses@isb-sib.ch](mailto:licenses@isb-sib.ch)).  
DR EMBL; BC002575; AA002575.1; ALT\_INIT.  
DR EMBL; BC037964; AA037964.1; -.  
DR EMBL; AB040950; BA040950.1; -.  
DR Genew; HGNC:17210; DHX37.  
DR InterPro; IPR001410; DEAD.  
DR InterPro; IPR001650; Helicase\_C.  
DR Pfam; PF00408; HA2; 1.  
DR Pfam; PF00271; helicase\_C; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR SMART; SM00490; HELIC; 1.  
DR PROSITE; PS00690; DEAD ATP HELICASE; FALSE\_NEG.  
DR KMW; KMW00000; Helicase; ATP-binding.  
FT NP BIND 275 282 ATP (POTENTIAL).  
FT SITE 372 375 DEAD BOX.  
FT CONFLICT 869 869 S -> G (IN REF. 1; AA002575).  
FT CONFLICT 898 898 A -> S (IN REF. 2).  
FT CONFLICT 1081 1081 R -> Q (IN REF. 1; AA002575).  
FT CONFLICT 1130 1157 YLLAEYCEWLPQAMHPDIKAMPPTVH -> CEFDQGGGVGV  
FT GVDPRGSLRQGLCALCTVSPGLAEGSGPTANGQLEAT (IN  
FT REF. 2).  
SQ SEQUENCE 1157 AA; 129544 MW; 49332175221B30C5 CRC64;  
Query Match 0.7%; Score 8; DB 1; Length 1157;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 380 VLLGVGCA 387  
DB 857 VLLGVGCA 864

RESULT 38  
ITAD HUMAN STANDARD; PRT; 1162 AA.  
ID Q13349; Q15575; Q15576;  
AC Q13349; Q15575; Q15576;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).  
GN ITGAD.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RX MEDLINE=96111956; PubMed=8777714;  
RA Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,  
RA Staunton D.E., Gallatin W.M.;  
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-3.";  
RL J. Biol. Chem. 275:8959-8969(2000).  
RN [3]  
RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.  
RX MEDLINE=96252336; PubMed=8666289;  
RA Wong D.A., Davis E.M., Lebeau M., Springer T.A.;  
RT "Cloning and chromosomal localization of a novel gene-encoding a human  
RT beta 2-integrin alpha subunit.";  
RL Gene 171:291-294(1996).  
RN [4]  
RP INTERACTION WITH VCAM1.  
RX MEDLINE=99059842; PubMed=9841932;  
RA Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,  
RA Hoffman P.A., Staunton D.E., Bochner B.S.;  
RT "alpha2beta2 integrin is expressed on human eosinophils and functions  
RT as an alternative ligand for vascular cell adhesion molecule 1  
RT (VCAM-1).";  
RL J. Exp. Med. 188:2187-2191(1998).  
RN [5]  
RP INTERACTION WITH VCAM1.  
RX MEDLINE=99370002; PubMed=10438935;  
RA Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,  
RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;  
RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a  
RT binding interface between I domain and VCAM-1.";  
RL J. Immunol. 163:1984-1990(1999).  
CC -!- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND  
CC VCAM1. MAY PLAY A ROLE IN THE ATHEROSCLEROTIC PROCESS SUCH AS  
CC CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-  
CC BORNE PATHOGENS, PARTICULATE MATTER, AND SENESECENT ERYTHROCYTES  
CC FROM THE BLOOD.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D  
CC ASSOCIATES WITH BETA-2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL  
CC LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON  
CC TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN  
CC ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.  
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSET) IS A VWFA DOMAIN. INTEGRINS  
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch).

EMBL;	U37028;	AAB38547.1;	-	
DR	EMBL;	U40274;	AAB60634.1;	-
DR	EMBL;	U40275;	AAB60635.1;	-
DR	EMBL;	U40276;	AAB60636.1;	-
DR	EMBL;	U40277;	AAB60637.1;	-
DR	EMBL;	U40278;	AAB60638.1;	-
DR	EMBL;	U40279;	AAB60639.1;	-
DR	EMBL;	U40278;	AAB60638.1;	JOINED.
DR	EMBL;	AF187881;	AAF62875.1;	-
DR	HSSP;	Fl1215;	1A8X.	
DR	Genew;	HGNC:	6146;	ITGAD.
DR	MIM;	602453;	-	
DR	GO;	GO:0008305;	C:integrin complex;	TAS.
DR	GO;	GO:0004895;	F:cell_adhesion receptor activity;	TAS.
DR	GO;	GO:0016337;	P:cell-cell adhesion;	NAS.
DR	GO;	GO:0007160;	P:cell-matrix adhesion;	NAS.
DR	GO;	GO:0006955;	P:immune response;	NAS.
DR	InterPro;	IPR000413;	Integrin_alpha.	
DR	InterPro;	IPR002035;	VWF_A.	
DR	Pfam;	PF01839;	FG-GAP; 3.	
DR	Pfam;	PF00357;	integrin_A; 1.	
DR	Pfam;	PF00032;	vwa; 1.	
DR	PRINTS;	PRO1185;	INTEGRINA.	
DR	PRINTS;	PRO0453;	VWFADOMAIN.	
DR	SMART;	SM00191;	Int_alpha; 4.	
DR	SMART;	SM00327;	VWA; 1.	
DR	PROSITE;	PS00242;	INTEGRIN_ALPHA; 1.	
DR	PROSITE;	PS50234;	VWFA; 1.	
DR	PROSITE;	PS50234;	VWFA; 1.	
KW	Integrin;	Cell adhesion;	Receptor;	Glycoprotein; Transmembrane;
KW	Signal;	Repeat;	Calcium;	
KW	Magnesium.			
FT	SIGNAL	1	17	POTENTIAL.
FT	CHAIN	18	1162	INTEGRIN ALPHA-D.
FT	DOMAIN	18	1100	EXTRACELLULAR (POTENTIAL).
FT	TRANSKEM	1101	1124	POTENTIAL.
FT	DOMAIN	1125	1162	CYTOPLASMIC (POTENTIAL).
FT	REPEAT	32	85	FG-GAP 1.
FT	REPEAT	?	?	FG-GAP 2.
FT	DOMAIN	150	332	VWFA.
FT	REPEAT	350	400	FG-GAP 3.
FT	REPEAT	401	452	FG-GAP 4.
FT	REPEAT	454	516	FG-GAP 5.
FT	REPEAT	518	576	FG-GAP 6.
FT	REPEAT	581	633	FG-GAP 7.
FT	CA_BIND	465	473	POTENTIAL.
FT	CA_BIND	530	538	POTENTIAL.
FT	CA_BIND	593	601	POTENTIAL.
FT	SITE	1127	1131	GFPR MOTIF.
FT	DISULFID	67	74	BY SIMILARITY.
FT	DISULFID	106	124	BY SIMILARITY.
FT	DISULFID	655	710	BY SIMILARITY.
FT	DISULFID	769	775	BY SIMILARITY.
FT	DISULFID	846	861	BY SIMILARITY.
FT	DISULFID	994	1018	BY SIMILARITY.
FT	DISULFID	1023	1028	BY SIMILARITY.
FT	CARBOHYD	59	59	N-LINKED (GLCNAC. .)
FT	CARBOHYD	87	87	N-LINKED (GLCNAC. .)
FT	CARBOHYD	99	99	N-LINKED (GLCNAC. .)
FT	CARBOHYD	391	391	N-LINKED (GLCNAC. .)
FT	CARBOHYD	691	691	N-LINKED (GLCNAC. .)
FT	CARBOHYD	733	733	N-LINKED (GLCNAC. .)
FT	CARBOHYD	873	873	N-LINKED (GLCNAC. .)
FT	CARBOHYD	957	957	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1046	1046	N-LINKED (GLCNAC. .)
FT	CONFLICT	500	500	MISSING (IN REF. 2).
FT	CONFLICT	515	518	CHPW -> ATP (IN REF. 2).
FT	CONFLICT	825	825	L -> V (IN REF. 2).

```

FT CONFLICT 984 984 V -> A (1N REP. 2).
SQ SEQUENCE 1162 AA; 126985 MM; F296A1A3545D77D CRC64;

Query Match 0.7%; Score 8; DB 1; Length 1162;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
DB 47 LVVGAPLE 54

RESULT 39
ITAX HUMAN
ID ITAX HUMAN STANDARD; PRT; 1163 AA.
AC AC P20702;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-X precursor (leukocyte adhesion glycoprotein p150,95
DE alpha chain) [leukocyte adhesion receptor p150,95] (CD11c) (Leu M5).
GN ITGAX OR CD11c.
OS Homo sapiens [Human].
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166645; PubMed=3327687;
RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.; of
RT "cDNA cloning and complete primary structure of the alpha subunit of
RT a leukocyte adhesion glycoprotein, p150,95."
RL EMBO J. 6:4023-4028(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90151906; PubMed=2303426;
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RT "Genomic structure of an integrin alpha subunit, the leukocyte
RT p150,95 molecule."
RL J. Biol. Chem. 265:2782-2788(1990).
RN [3]
RP ERRATUM.
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RL J. Biol. Chem. 265:12750-12751(1990).
RN [4]
RP SEQUENCE OF 20-43.
RX MEDLINE=87167596; PubMed=3549901;
RA Miller L.J., Wiebe M., Springer T.A.;
RT "Purification and alpha subunit N-terminal sequences of human Mac-1
RT and p150,95 leukocyte adhesion proteins."
RL J. Immunol. 138:2381-2383(1987).
CC -1- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT
CC RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL
CC INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY
CC IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X
CC ASSOCIATES WITH BETA-2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFPA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
CC -1- SIMILARITY: Contains 1 VFPA domain.
CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC -----

```

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

QY	1151	QGLMLAL	1158
DB	1115	QGLMLAL	1122
RESULT 40			
TSPI_XENLA		STANDARD;	PRT; 1173 AA.
ID	TSPI_XENLA		
AC	P35448;		
DT	01-JUN-1994	(Rel. 29, Created)	
DT	01-JUN-1994	(Rel. 29, Last sequence update)	
DT	28-FEB-2003	(Rel. 41, Last annotation update)	
DE	Thrombospondin 1 precursor.		
GN	THBS1 OR TSPI.		
OS	Xenopus laevis (African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;		
OC	Xenopodidae; Xenopus.		
OX	NCBI_TaxID=8355;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;		
RL	Submitted (XXX-1993) to the EMBL/GenBank/DBJ databases.		
CC	FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and cell-to-matrix interactions. Can bind to fibronectin, laminin, type V collagen and integrins alpha-V/beta-1, alpha-V/beta-3 and alpha-1Ib/beta-3 (by similarity).		
CC	SUBUNIT: Homotrimer; disulfide-linked.		
CC	SIMILARITY: Belongs to the thrombospondin family.		
CC	SIMILARITY: Contains 1 VWFC domain.		
CC	SIMILARITY: Contains 3 EGF-like domains.		
CC	SIMILARITY: Contains 3 TSP type-1 domains.		
CC	SIMILARITY: Contains 7 TSP type-3 domains.		
CC	SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcement">http://www.isb-sib.ch/announcement</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	EMBL; L04278; -; NOT_ANNOTATED_CDS.		
DR	HSSP; P00740; 1EDM.		
DR	InterPro; IP001881; EGF_Ca.		
DR	InterPro; IP006209; EGF-like.		
DR	InterPro; IP008210; IBGF.		
DR	InterPro; IP008884; TSPI.		
DR	InterPro; IP008085; TSP_1.		
DR	InterPro; IP003367; tsp_3.		
DR	InterPro; IP008859; TSPC.		
DR	InterPro; IP003129; TSPN.		
DR	InterPro; IP001007; VWFC_C.		
DR	Pfam; PF00008; EGF; 2.		
DR	Pfam; PF00090; tsp_1; 3.		
DR	Pfam; PF02412; tsp_3; 13.		
DR	Pfam; PF05735; TSPC; 1.		
DR	Pfam; PF02210; TSPN; 1.		
DR	Pfam; PF00093; vwc; 1.		
DR	PRINTS; PR01705; TSP1REPEAT.		
DR	SMART; SM00181; EGF; 2.		
DR	SMART; SM00209; TSP1; 3.		
DR	SMART; SM00210; TSPN; 1.		
DR	SMART; SM00214; VWC; 1.		
DR	PROSITE; PS00022; EGF_1; FALSE_NEG.		
DR	PROSITE; PS01186; EGF_2; 1.		
DR	PROSITE; PS00026; EGF_3; 2.		
DR	PROSITE; PS00092; TSP1; 3.		
DR	PROSITE; PS01208; VWFC_1; 1.		
DR	PROSITE; PS00184; VWFC_2; 1.		
KW	Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;		

```
Query Match      0.7%; Score 8; DB 1; Length 1163;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels
```

EGF-like domain; Signal.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
--------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--



```

RESULT 42
PMPC_CHLMU
ID PMPC_CHLMU STANDARD; PRT; 1460 AA.
AC Q9P9J1
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmpc precursor (Polymorphic membrane
DE protein C).
DE PMPC OR TC0695.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
[1]
SEQUENCE FROM N.A.
RN STRAIN=MoPn / Ni9g;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., Deboy K., Kolonay J., McClarty G., Salzberg S.L.,
RA Bisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39";
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -!- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
CC (Potential).
CC -!- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB002338; AAP39511.1; --
CC PIR; D81675; D81675.
CC TIGR; TC0695; --
CC InterPro; IPR006315; Autotransport.
CC InterPro; IPR003368; Chlamydia PMP.
CC Pfam; PF02415; Chlamydia PMP; 2.
CC TIGRFAMs; TIGR01414; autotrans bar1; 1.
CC TIGRFAMs; TIGR01376; POMP repeat; 8.
CC Outer membrane; signal; Multigene family; Complete proteome.
CC SIGNAL 1 20 POTENTIAL.
CC CHAIN 21 1460 PROBABLE OUTER MEMBRANE PROTEIN PMPC.
CC SEQUENCE 1460 AA; 154781 MW; C1F033E7907AC3AC CRC64;
Query Match 0.7%; Score 8; DB 1; Length 1460;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 583 SILKTPKQ 590
DB 1311 SILKTPKQ 1318
[1]
RESULT 43
SET2_CABEL
ID SET2_CABEL STANDARD; PRT; 1507 AA.
AC Q182z1; Q95QU6; Q95QU7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein set-2.
DE SET-2 OR C26E6.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]

```

RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Fulton L.;  
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP REVISIONS, AND ALTERNATIVE SPLICING.  
RA Waterston R.;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
RX MEDLINE=21585681; PubMed=11729150;  
RA Xu L., Strone S.;  
RL "Depletion of a novel SET-domain protein enhances the sterility of  
mes-3 and mes-4 mutants of *Caenorhabditis elegans*.";  
RT Genetics 159:1019-1029(2001).  
CC -1- FUNCTION: Probably involved in chromatin modification and/or  
remodelling in meiotic germ cells. May act redundantly with mes-3  
and mes-4 proteins.  
CC -1- SUBCELLULAR LOCATION: Nuclear. Localized in mitotic and mid-late-  
stage meiotic nuclei but is undetectable in early pachytene  
nuclei.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=a; Synonyms=L;  
CC IsoId=Q18221-1; Sequence=Displayed;  
CC Name=b; Synonyms=S;  
CC IsoId=Q18221-2; Sequence=VSP\_007217, VSP\_007218;  
CC -1- TISSUE SPECIFICITY: Expressed in all cells of embryo. In L1 larva,  
it is predominantly expressed in Z2 and Z3 primordial germ cells.  
CC In adult is predominantly expressed in the germline.  
CC -1- DEVELOPMENTAL STAGE: Expressed throughout embryogenesis.  
CC -1- SIMILARITY: Contains 1 post-SET domain.  
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.  
CC -1- SIMILARITY: Contains 1 SET domain.  
CC -1- CAUTION: Ref.1 (AAK67215) sequence differs from that shown due to  
erroneous gene model prediction.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U13875; AAA21163.1; -;  
CC EMBL; U13875; AAK67214.1; -;  
CC EMBL; U13875; AAK67215.1; ALT\_SEQ.  
CC PIR; A88445; A88445  
CC Wormpep; C2656.9a; CE27735.  
CC Wormpep; C2656.9b; CEU1158.  
CC InterPro; IPR003616; PostSET.  
CC InterPro; IPR000504; RNA\_rec\_mot.  
CC InterPro; IPR001214; SET.  
CC Pfam; PF00856; SET; 1.  
CC SMART; SM00508; PostSET; 1.  
CC SMART; SM00360; RRM; 1.  
CC SMART; SM00317; SET; 1.  
CC PROSITE; PS50868; POST SET; 1.  
CC PROSITE; PS50102; RRM; FALSE NEG.  
CC PROSITE; PS00030; RRM\_RNP\_1; FALSE\_NEG.  
CC PROSITE; PS50280; SET; 1.  
KW RNA-binding; Nuclear protein; Alternative splicing.  
FT DOMAIN 128 199 RNA-BINDING (RRM) (POTENTIAL).  
FT DOMAIN 1366 1489 SET.  
FT DOMAIN 1491 1507 POST-SET.  
FT DOMAIN 296 354 PRO-RICH.  
FT DOMAIN 554 664 PRO-RICH.  
FT DOMAIN 870 1011 SER-RICH.  
FT VARSPPLIC 1 768 Missing (in isoform b).  
FT /FTID=VSP\_007217.  
FT MDSLRKVAEDIRQIMRQCFAALDKHLKAIADKKKK  
FT EREKRAQEAEPNSHLIADMM -> MYNNSAPYLNHSSLN

PT TVRKQVTVRRVLPSPPPPPPPPSLPSPSPSVKVPYIPQR  
FT VYRSINS (in isoform b).  
FT /FTID=VSP\_007218.  
SQ SEQUENCE 1507 AA; 171681 MW; E7D9689DA720C34A CRC64;  
Query Match 0.7%; Score 8; DB 1; Length 1507;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 430 TSVVSSRQ 437  
Db 1002 TSVVSSRQ 1009  
RESULT 44  
ID RPA1 YEAST STANDARD; PRT: 1664 AA.  
AC P10964; Q99330;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE DNA-directed RNA polymerase I 190 kDa polypeptide (EC 2.7.7.6) (A190).  
GN RPA1 OR RPA190 OR RRN1 OR YOR341W OR O6276.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88139333; PubMed=2830265;  
RA Memet S., Gouy M., Marck C., Sentenac A., Buhler J.-M.;  
RT "RPA190, the gene coding for the largest subunit of yeast RNA  
polymerase A";  
RL J. Biol. Chem. 263:2830-2839(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / FY73;  
RX MEDLINE=97103776; PubMed=8948102;  
RA Purnelle B., Goffeau A.;  
RT "Nucleotide sequence analysis of a 40 kb segment on the right arm of  
yeast chromosome XV reveals 18 open reading frames including a new  
pyruvate kinase and three homologues to chromosome I genes.";  
RT Yeast 12:1475-1481(1996).  
RL Yeast 12:1475-1481(1996).  
CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription  
of DNA into RNA using the four ribonucleoside triphosphates as  
substrates. RNA polymerase I is essentially used to transcribe  
ribosomal DNA units.  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
{RNA}(N)  
CC -1- SUBUNIT: RNA polymerase I consists of 14 different subunits. This  
subunit is the largest component of RNA polymerase I.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases were  
found in eukaryotic nuclei: polymerase I for the ribosomal RNA  
precursor, polymerase II for the mRNA precursor, and polymerase  
III for 5S and 5.8S rRNA genes.  
CC -1- SIMILARITY: Belongs to the RNA polymerase beta' chain family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; J03530; AAA34890.1; -;  
CC EMBL; X95720; CAA65029.1; -;  
CC EMBL; Z75249; CAA99665.1; -;  
CC PIR; S67250; S67250.  
CC GeneOnline; 143929;  
CC InterPro; IPR000722; RNA\_pol\_A.

```

DR InterPro; IPR007080; RNA_pol_Rpb1_1.
DR InterPro; IPR007086; RNA_pol_Rpb1_3.
DR InterPro; IPR007083; RNA_pol_Rpb1_4.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR InterPro; IPR006592; RNA_pol_A_N.
DR Pfam; PF04997; RNA_pol_Rpb1_1; 1.
DR Pfam; PF00623; RNA_pol_Rpb1_2; 1.
DR Pfam; PF04983; RNA_pol_Rpb1_3; 1.
DR Pfam; PF05000; RNA_pol_Rpb1_4; 1.
DR Pfam; PF04998; RNA_pol_Rpb1_5; 1.
DR SMART; SMO0663; RPOLA_N; 1.
KW Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
KW Zinc-finger; Nuclear protein.
FT ZN_FING 62 78 C2H2-TYPE (POTENTIAL).
FT CONFLICT 158 158 N -> T (IN REF. 1).
SQ SEQUENCE 1664 AA; 186431 MW; DF6SAVAA459DSE6D CRC64;

Query Match 0.7%; Score 8; DB 1; Length 1664;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 481 GSEITSVD 488
DB 8 GSEITSVD 15

RESULT 45
RRPL_BUNYV
ID _RRPL_BUNYV STANDARD; PRT; 2238 AA.
AC P20470;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE RNA-directed RNA polymerase (EC 2.7.7.48) (L protein).
GN L.
OS Bunyamwera virus.
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Orthobunyavirus.
OX NCBI_TaxID=35304;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90085791; PubMed=2596023;
RA Elliott R.M.;
RT "Nucleotide sequence analysis of the large (L) genomic RNA segment of Bunyamwera virus, the prototype of the family Bunyaviridae.";
RL Virology 173:426-436(1989).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA}(N).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X14383; CAA32553.1; -.
CC PIR; A33744; RRVUBY.
DR InterPro; IPR007322; Bunya_RdRp.
DR InterPro; IPR007099; RNA_pol_NSVir.
DR Pfam; PF04196; Bunya_RdRp; 1.
KW Transferrase; Nucleotidyltransferase; RNA-directed RNA polymerase.
SQ SEQUENCE 2238 AA; 258668 MW; 1ED00AB156BAC8DA CRC64;

Query Match 0.7%; Score 8; DB 1; Length 2238;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 FLINILKK 191
DB 1333 FLINILKK 1340

```

```

RESULT 46
LSTL_MOUSE
ID LSTL_MOUSE STANDARD; PRT; 95 AA.
AC O08843; O08844; O92IH3; Q9ZIR0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Leukocyte specific transcript 1 protein (B144 protein).
GN LSTL OR B144.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE-B-cell;
RX MEDLINE=88031493; PubMed=3117682;
RA Tsuge I., Shen F.-W., Steinmetz M., Boyse E.A.;
RT "A gene in the H-2S.H-2D interval of the major histocompatibility
RT complex which is transcribed in B cells and macrophages.";
RL Immunogenetics 26:378-380(1987).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN-BALB/c; TISSUE-Macrophage;
RX MEDLINE=98035883; PubMed=9367684;
RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
RA Weiss E.H.;
RT "Complex expression pattern of the TNF region gene LST1 through
RT differential regulation, initiation, and alternative splicing.";
RL Genomics 45:591-600(1997).
RN [3]
RP SEQUENCE FROM N.A.
RA Rowen L., Qin S., Madan A., Abbasi N., James R., Dickhoff R.,
RA Shaffer T., Ratcliffe A., Loretz C., Lasky S., Hood L.;
RT "Sequence of the mouse major histocompatibility class III region.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE=21372017; PubMed=11478849;
RA Raghunathan A., Sivakamasundari R., Wolenski J., Poddar R.,
RA Weissman S.M.;
RT "Functional analysis of B144/LST1: a gene in the tumor necrosis
RT factor cluster that induces formation of long filopodia in eukaryotic
RT cells.";
RL Exp. Cell Res. 268:230-244(2001).
CC -!- FUNCTION: Possible role in modulating immune responses. Has an
CC inhibitory effect on lymphocyte proliferation. Induces
CC morphological changes including production of filopodia and
CC microspikes when overexpressed in a variety of cell types and may
CC be involved in dendritic cell maturation.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Also detected in
CC a perinuclear region corresponding to the localization of the
CC Golgi apparatus and throughout the cytoplasm (By
CC similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=ml7r;
CC IsoId=O08843-1; Sequence=Displayed;
CC Name=2; Synonyms=ml21r;
CC IsoId=O08843-2; Sequence=VSP_050588;
CC -!- TISSUE SPECIFICITY: Expressed in spleen and at lower levels in
CC thymus and liver.
CC -!- SIMILARITY: Belongs to the LST1 family.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 58.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC "Cat allergen 1: Biochemical, antigenic, and allergenic properties";  
CC J. Allergy Clin. Immunol. 74:147-153 (1984).  
CC -!- SUBUNIT: Heterotrimer composed of two non-covalently linked  
CC disulfide-linked heterodimer of chains 1 and 2.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Comment=Experimental confirmation may be lacking for some  
CC isoforms;  
CC Name=1; Synonyms=CH2L;  
CC IsoId=P30440-1; Sequence=Displayed;  
CC Name=2; Synonyms=CH2S;  
CC IsoId=P30440-2; Sequence=VSP\_004249;  
CC Name=3; Synonyms=CH2ST, Truncated;  
CC IsoId=P30440-3; Sequence=VSP\_004249;  
CC -!- TISSUE SPECIFICITY: The long form is preferentially expressed in  
CC the salivary gland, while the short form is preferentially  
CC expressed in the skin.  
CC -!- ALLERGEN: Causes an allergic reaction in human. Major allergen  
CC produced by the domestic cat.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M77341; AAC41616.1; -;  
CC EMBL; X62478; CAA44345.1; -;  
CC PIR; B53283; B53283.  
CC PIR; C56413; C56413.  
CC PIR; J01127; J01127.  
CC InterPro: IPR006038; Uteroglobin supf.  
CC Allergen; Glycoprotein; Signal; Polymorphism; Alternative splicing.  
CC SIGNAL 1 17  
CC CHAIN 18 109  
CC CARBOHYD 50 50  
CC VARSPPLIC 82 109  
CC  
CC TTISSKDCMGSAVNTVDLKLTLGR -> PSTNIWVK  
CC QFRTF (in isoform 3).  
CC /FTID-VSP\_004248  
CC TTISSKDC -> IAINCY (in isoform 2).  
CC /FTID-VSP\_004249  
CC I -> L (IN CH2LV).  
CC I -> V (IN CH2SV).  
CC RV -> KP (IN CH2SV).  
CC M -> T (IN CH2LV).  
CC Q -> E (IN CH2SV).  
CC N -> K (IN CH2SV).  
CC C -> F (IN REF. 3).  
CC F -> T (IN REF. 3).  
CC SEQUENCE 109 AA; 11854 MW; 857FB9CD76036CB9 CRC64;  
CC  
CC Query Match 0.6%; Score 7; DB 1; Length 109;  
CC Best Local Similarity 100.0%; Pred. No. 53;  
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC QY 1150 LGGLLLLL 1156  
CC |||||  
CC Db 24 LGGLLLLL 30

RESULT 47  
ID FEL2\_FELCA STANDARD; PRT; 109 AA.

AC P30440;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Major allergen I polypeptide chain 2 precursor (Allergen Fel d 1-B)  
DE (Fel d 1-B) (Allergen Cat-I) (AG4) (Fdi).  
GN CH2.  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
OX NCBI\_TaxID=9685;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-100.  
RP MEDLINE=92052157; PubMed=1946388;  
RA Morgenstern J.P., Griffith I.J., Brauer A.W., Rogers B.L.,  
RA Bond J.F., Chapman M.D., Kuo M.-C.;  
RT "Amino acid sequence of Fel d1, the major allergen of the domestic  
RT cat: protein sequence analysis and cDNA cloning.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:9690-9694 (1991).  
RN [2]  
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND VARIANTS.  
RP TISSUE=Liver;  
RX MEDLINE=92241678; PubMed=1572548;  
RA Griffith I.J., Craig S., Pollock J., Yu X.-B., Morgenstern J.P.,  
RA Rogers B.L.;  
RT "Expression and genomic structure of the genes encoding Fd1, the  
RT major allergen from the domestic cat.";  
RL Gene 113:263-268 (1992).  
RN [3]  
RP SEQUENCE OF 18-37, AND CHARACTERIZATION.  
RX MEDLINE=91287714; PubMed=1712068;  
RA Duffort O.A., Carreira J., Nitti G., Polo F., Lombardero M.;  
RT "Studies on the biochemical structure of the major cat allergen Felis  
RT domesticus I.";  
RL Mol. Immunol. 28:301-309 (1991).  
RN [4]  
RP CHARACTERIZATION.  
RX MEDLINE=84265679; PubMed=6747135;  
RA Leiterman K., Ohman J.L. Jr.;

Query Match 0.6%; Score 7; DB 1; Length 109;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1157 ALLVLAL 1163  
Db 4 ALLVLAL 10

RESULT 48  
HDEA\_ECOLI  
ID HDEA\_ECOLI STANDARD; PRT; 110 AA.

AC P26604;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Protein hdeA precursor (10K-S protein).  
GN HDEA OR B3510 OR Z4922 OR ECS4390 OR SF3544 OR S4223.  
OS Escherichia coli,  
OS Escherichia coli O157:H7, and

OS *Shigella flexneri*.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562, 83334, 623;  
RN [1]  
RN SEQUENCE FROM N.A., AND SEQUENCE OF 22-40.  
RP SPECIES=E.coli; STRAIN=K12;  
RC MEDLINE=94316500; PubMed=8041620;  
RX Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;  
RA "Analysis of the Escherichia coli genome. V. DNA sequence of the  
RT region from 76.0 to 81.5 minutes.";  
RL Nucleic Acids Res. 22:2576-2586(1994).  
RN [3]  
RN SEQUENCE FROM N.A.  
RP SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 7009527;  
RC MEDLINE=21074935; PubMed=11206551;  
RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
RL Nature 409:523-533 (2001).  
RN [14]  
RN SEQUENCE FROM N.A.  
RP SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;  
RC MEDLINE=21156231; PubMed=11258796;  
RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Iehii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
RL DNA Res. 8:11-22(2001).  
RN [5]  
RN GENE NAME.  
RP SPECIES=E.coli;  
RC MEDLINE=94064579; PubMed=8244952;  
RX Yoshida T., Ueguchi C., Mizuno T.;  
RA "Physical map location of a set of Escherichia coli genes (hde) whose  
RT expression is affected by the nucleoid protein H-NS.";  
RL J. Bacteriol. 175:7747-7748(1993).  
RN [6]  
RN SEQUENCE OF 22-41.  
RP SPECIES=E.coli; STRAIN=K12 / W3110;  
RC Pasquali C., Sanchez J.-C., Ravier F., Golaz O., Hughes G.J.,  
RA Frutiger S., Paquet N., Wilkins M., Appel R.D., Bairoch A.,  
RA Hochstrasser D.F.;  
RL Submitted (SEP-1994) to Swiss-Prot.  
RN [7]  
RN SEQUENCE OF 22-33.  
RP SPECIES=E.coli; STRAIN=K12 / ENG2;  
RX MEDLINE=97443975; PubMed=9298646;  
RA Link A.J., Robison K., Church G.M.;  
RT "Comparing the predicted and observed properties of proteins encoded  
RT in the genome of Escherichia coli K-12.";  
RL Electrophoresis 18:1259-1313(1997).  
RN [8]  
RN SEQUENCE OF 22-31.  
RP SPECIES=E.coli; STRAIN=K12;  
RC MEDLINE=99085675; PubMed=9868784;  
RX Wasinger V.C., Humphrey-Smith I.;  
RA "Small genes/gene-products in Escherichia coli K-12.";  
RT

RL RN SEQUENCE FROM N.A.  
RP SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;  
RX MEDLINE=2227406; PubMed=12384590;  
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
RA Sun L., Yue Y., Zhao A., Cao Y., Zhu J., Kan B., Ding K., Chen S.,  
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
RA Yu J.;  
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity  
RT through comparison with genomes of Escherichia coli K12 and O157.";  
RL Nucleic Acids Res. 30:4432-4441(2002).  
RN [10]  
RN SEQUENCE FROM N.A.  
RP SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;  
RX MEDLINE=22590274; PubMed=12704152;  
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,  
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,  
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,  
RA Schwartz D.C., Blattner F.R.;  
RT "Complete genome sequence and comparative genomics of Shigella  
RT flexneri serotype 2a strain 2457T.";  
RL Infect. Immun. 71:2775-2786(2003).  
RN [11]  
RN MASS SPECTROMETRY.  
RP SPECIES=E.coli; STRAIN=ATCC 15597;  
RC MEDLINE=22067786; PubMed=12071744;  
RX Reid G.E., Shang H., Hogan J.M., Lee G.U., McLuckey S.A.;  
RA "Gas-phase concentration, purification, and identification of whole  
RT proteins from complex mixtures.";  
RL J. Am. Chem. Soc. 124:7353-7362(2002).  
RN [12]  
RN X-RAY CRYSTALLOGRAPHY.  
RP SPECIES=E.coli;  
RC MEDLINE=98400493; PubMed=9731767;  
RX Yang F., Gustafson K.R., Boyd M.R., Wlodawer A.;  
RT "Crystal structure of Escherichia coli HdeA.";  
RL Nat. Struct. Biol. 5:763-764(1998).  
RN [13]  
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RP SPECIES=E.coli;  
RC MEDLINE=20090957; PubMed=10623350;  
RX Gajiwala K.S., Burley S.K.;  
RA "HDEA, a periplasmic protein that supports acid resistance in  
RT pathogenic enteric bacteria.";  
RL J. Mol. Biol. 295:605-612(2000).  
CC -1- SUBCELLULAR LOCATION: Periplasmic (Potential).  
CC -1- MASS SPECTROMETRY: MW=9742; METHOD=Electrospray; RANGE=22-110.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC  
CC EMBL: D11109; BAA01883.1; -;  
CC EMBL: U03039; AAE18486.1; -;  
CC EMBL: AE000427; AAC76535.1; -;  
CC EMBL: AE005576; AAG58851.1; -;  
CC EMBL: AP002565; BAB37813.1; -;  
CC EMBL: AE015362; AANA4999.1; -;  
CC EMBL: AE016992; AAP19187.1; -;  
CC PIR: F91177; F91177;  
CC PIR: G86023; G86023;  
CC PIR: S30268; S30268;  
CC PDB: 1BG8; 16-SEP-98.  
CC PDB: 1DJ8; 10-DEC-99.  
CC DR SWISS-2DPAGE; P26604; COLI.  
CC DR EcoGene; EG11398; hdeA.  
CC Periplasmic; Signal; 3D-structure; Complete proteome.

```

FT SIGNAL 1 21
FT CHAIN 22 110
FT DISULFID 39 87
FT HELIX 34 36
FT STRAND 38 38
FT HELIX 39 43
FT TURN 44 44
FT HELIX 47 49
FT HELIX 50 60
FT TURN 61 64
FT HELIX 66 68
FT HELIX 73 88
FT TURN 89 89
FT TURN 91 92
FT STRAND 94 94
FT HELIX 95 107
SQ SEQUENCE 110 AA; 11858 MW; 063262C4863FA2B9 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1150 LGGLLLL 1156
Db 9 LGGLLLL 15

RESULT 49
INS1_RAT
ID INS1_RAT STANDARD; PRT; 110 AA.
AC P01322;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DB Insulin 1 precursor.
GN INS1 OR INS-1
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=80045034; PubMed=498283;
RA Cordell B., Bell G.I., Fischer E., Denoto F.M., Ullrich A.,
Pictet R.L., Rutter W.J., Goodman H.M.;
RT "Isolation and characterization of a cloned rat insulin gene.";
RL Cell 18:533-543(1979).
[2]
SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; Tissue=Liver;
RX MEDLINE=80045035; PubMed=498284;
RA Lomedico P., Rosenthal N., Estratiadis A., Gilbert W., Kolodner R.,
Tizard R.;
RT "The structure and evolution of the two nonallelic rat preproinsulin
genes.";
RL Cell 18:545-558(1979).
[3]
SEQUENCE FROM N.A.
RX MEDLINE=80240379; PubMed=6249167;
RA Lomedico P.T., Rosenthal N., Kolodner R., Estratiadis A.,
Gilbert W.;
RT "The structure of rat preproinsulin genes.";
RL Ann. N.Y. Acad. Sci. 343:425-432(1980).
[4]
SEQUENCE OF 25-54 AND 90-110.
RX MEDLINE=70067613; PubMed=4311938;
RA Steiner D.F., Clark J.L., Nolan C., Rubenstein A.H., Margoliash E.,
Aten B., Oyer P.E.;
RT "Proinsulin and the biosynthesis of insulin.";
RL Recent Prog. Horm. Res. 25:207-282(1969).
[5]
SEQUENCE OF 57-87.
RX MEDLINE=73061498; PubMed=4640931;

```

```

RA Tager H.S., Steiner D.F.;
RT "Primary structures of the proinsulin connecting peptides of the rat
and the horse.";
RL J. Biol. Chem. 247:7936-7940(1972).
[6]
RP SEQUENCE OF 57-87, AND REVISIONS.
RX MEDLINE=72177385; PubMed=4554104;
RA Markussen J., Sundby F.;
RT "Rat-proinsulin C-peptides. Amino-acid sequences.";
RL Eur. J. Biochem. 25:153-162(1972).
CC -!- FUNCTION: Insulin decreases blood glucose concentration. It
increases cell permeability to monosaccharides, amino acids and
fatty acids. It accelerates glycolysis, the pentose phosphate
cycle, and glycogen synthesis in liver.
CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
disulfide bonds.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the insulin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
DR EMBL; V01242; CA24559.1; -
DR EMBL; J00747; AAA41442.1; -
DR EMBL; M25584; AAA41439.1; -
DR PIR; A90788; IPRTI.
DR HSSP; P01308; 1A7F.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IIGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism; Signal; Multigene family.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN 1 B CHAIN.
FT PROPEP 57 87 INSULIN 1 C PEPTIDE.
FT CHAIN 90 110 INSULIN 1 A CHAIN.
FT DISULFID 31 96 INTERCHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100
SQ SEQUENCE 110 AA; 12420 MW; 51D606DA54AE3533 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1155 LLALLVL 1161
Db 10 LLALLVL 16

RESULT 50
INS_RABIT
ID INS_RABIT STANDARD; PRT; 110 AA.
AC P01311;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Insulin precursor.
GN INS.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; Tissue=Pancreas;
RX MEDLINE=94179230; PubMed=8132571;

```

RA Devaskar S.U., Giddings S.J., Rajakumar P.A., Carnaghi L.R.,  
 RA Menon R.K., Zahm D.S.,  
 RT "Insulin gene expression and insulin synthesis in mammalian neuronal  
 cells.", Chem. 269:8445-8454(1994).  
 RL J. Biol. Chem. 269:8445-8454(1994).  
 RN [2]  
 RP SEQUENCE OF 25-54 AND 90-110.  
 RX MEDLINE=66160119; PubMed=5949593;  
 RA Smith L.P.;  
 RT "Species variation in the amino acid sequence of insulin.";  
 RL Am. J. Med. 40:662-666(1966).  
 RN [3]  
 RP SEQUENCE OF 56-110 FROM N.A.  
 RA Giddings S.J., Carnaghi L.R., Devaskar S.U.;  
 RL Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Insulin decreases blood glucose concentration. It  
 CC increases cell permeability to monosaccharides, amino acids and  
 CC fatty acids. It accelerates glycolysis, the pentose phosphate  
 CC cycle, and glycogen synthesis in liver.  
 CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two  
 CC disulfide bonds.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the insulin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U03610; AA19033.1; -;  
 DR EMBL; M61153; AA17540.1; -;  
 DR PIR; A53438; INRS.  
 DR HSSP; P01308; ILYM.  
 DR InterPro; IPR004925; Ins/IGF/relax.  
 DR Pfam; PF00049; Insulin; 1.  
 DR PRINTS; PR00277; INSULINB.  
 DR SMART; SM00078; IIGF; 1.  
 DR PROSITE; PS00262; INSULIN; 1.  
 DR Insulin family; Hormone; Glucose metabolism; Signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 54 INSULIN B CHAIN.  
 FT PROPEP 57 87 C PEPTIDE.  
 FT CHAIN 90 110 INSULIN A CHAIN.  
 FT DISULFID 31 96 INTERCHAIN.  
 FT DISULFID 43 109 INTERCHAIN.  
 FT DISULFID 95 100 INTERCHAIN.  
 FT CONFLICT 83 83 E -> Y (IN REF. 3).  
 SQ SEQUENCE 110 AA; 11838 MW; 82D2975B85D77FA8 CRC64;  
 Query Match 0.6%; Score 7; DB 1; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1155 LLALLVL 1161  
 DB 10 LLALLVL 16  
 RESULT 51  
 IGF MYXGL  
 ID IGF MYXGL STANDARD; PRT; 139 AA.  
 AC P22518;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DE Insulin-like growth factor precursor (IGF) (Fragment).  
 OS Myxine glutinosa (Atlantic hagfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;  
 OC Myxiniidae; Myxinae; Myxine.  
 OX NCBI\_TaxID=7769;

[1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=91115860; PubMed=1989990;  
 RA Nagamatsu S., Chan S.J., Falkner S., Steiner D.F.;  
 RT "Evolution of the insulin gene superfamily. Sequence of a  
 preproinsulin-like growth factor cDNA from the Atlantic hagfish.";  
 RL J. Biol. Chem. 266:2397-2402(1991).  
 CC -!- FUNCTION: The insulin-like growth factors, isolated from plasma,  
 CC are structurally and functionally related to insulin but have a  
 CC much higher growth-promoting activity.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the insulin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; M57735; AAA49265.1; -;  
 DR PIR; A38612; A38612.  
 DR HSSP; P01344; IGF2.  
 DR InterPro; IPR004825; Ins/IGF/relax.  
 DR Pfam; PF00049; Insulin; 1.  
 DR PRINTS; PR00277; INSULINB.  
 DR SMART; SM00078; IIGF; 1.  
 DR PROSITE; PS00262; INSULIN; 1.  
 DR Insulin family; Mitogen; Growth factor; Signal.  
 FT NON TER 1 38  
 FT SIGNAL <1 38  
 FT CHAIN 39 139 INSULIN-LIKE GROWTH FACTOR.  
 FT DOMAIN 39 67 B.  
 FT DOMAIN 68 82 C.  
 FT DOMAIN 83 103 A.  
 FT DOMAIN 104 113 D.  
 FT DOMAIN 114 139 E.  
 SQ SEQUENCE 139 AA; 16087 MW; 2FC888C8D074FAC1 CRC64;  
 Query Match 0.6%; Score 7; DB 1; Length 139;  
 Best Local Similarity 100.0%; Pred. No. 77;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1153 LLLALL 1159  
 DB 28 LLLALL 34  
 RESULT 52  
 LAMP PETMA  
 ID LAMP PETMA STANDARD; PRT; 139 AA.  
 AC P33575; P33576;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DB Lamprin 0.9 precursor (Cartilage matrix protein).  
 OS Petromyzon marinus (Sea lamprey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;  
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.  
 OX NCBI\_TaxID=7757;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOPFORMS 0.9-10 AND 0.9-12), AND SEQUENCE OF  
 RP 20-44.  
 RC TISSUE=Cartilage;  
 RX MEDLINE=91123269; PubMed=7678258;  
 RA Robson P., Wright G.M., Sitarz E., Maiti A., Rawat M., Youson J.H.,  
 RA Keeley F.W.;  
 RT "Characterization of lamprin, an unusual matrix protein from lamprey  
 cartilage. Implications for evolution, structure, and assembly of  
 elastin and other fibrillar proteins.";  
 RL J. Biol. Chem. 268:1440-1447(1993).  
 CC -!- FUNCTION: Self-aggregating protein that makes part of the soluble

CC form of lamprin.  
 CC -1- SUBUNIT: The polymeric lamprin chains self-aggregate to form  
 CC fibers and have secondary structures particularly rich in beta-  
 CC sheets and in beta-turns.  
 CC -1- SUBCELLULAR LOCATION: Extracellular matrix.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=0.9-12;  
 CC IsoId=P33575-1; Sequence=Displayed;  
 CC Name=0.9-10;  
 CC IsoId=P33575-2; Sequence=VSP\_004302;  
 CC Note=No experimental confirmation available;  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC ENBL; L05925; AAA49269.1; --  
 CC ENBL; L05924; AAA49268.1; --  
 CC PIR; C45051; C45051.  
 CC Repeat; Connective tissue; Signal; Alternative splicing.  
 CC SIGNAL 1 19  
 CC CHAIN 20 139 LAMPRIN 0.9.  
 CC DOMAIN 42 110 8 X 5 AA APPROXIMATE REPEATS.  
 CC REPEAT 42 46 1.  
 CC REPEAT 47 51 2.  
 CC REPEAT 52 56 3.  
 CC REPEAT 57 61 4.  
 CC REPEAT 62 66 5.  
 CC REPEAT 67 71 6.  
 CC REPEAT 92 96 7.  
 CC REPEAT 106 110 8.  
 CC VARSPLIC 86 104 Missing (in isoform 0.9-10).  
 CC /Frida=VSP\_004302.  
 CC SEQUENCE 139 AA; 13257 MW; B248ABB7A6CEC7C3 CRC64;  
 CC -----  
 CC Query Match 0.63; Score 7; DB 1; Length 139;  
 CC Best Local Similarity 100.0%; Pred. No. 77;  
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC -----  
 CC Qy 1157 ALLVLAL 1163  
 CC | | | | |  
 CC 7 ALLVLAL 13  
 CC -----  
 CC RESULT 53  
 CC HEAD\_CHICK STANDARD; PRT; 141 AA.  
 CC AC P02001;  
 CC DT 21-JUL-1986 (Rel. 01, Created)  
 CC DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Hemoglobin alpha-D chain.  
 CC GN HBAD.  
 CC OS Gallus gallus (Chicken).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 CC OC Gallus.  
 CC CX NCBI TaxID=9031;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=White leghorn;  
 CC RX MEDLINE=92020223; PubMed=1656392;  
 CC RA Lewis W., Lee J.D., Dodgson J.B.;  
 CC RT "Adult chicken alpha-globin gene expression in transfected QT6 quail  
 CC RT cells: evidence for a negative regulatory element in the alpha D gene  
 CC RT region.";  
 CC RL Nucleic Acids Res. 19:5321-5329(1991).  
 CC RN [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=83161047; PubMed=6300093;  
 RA Dodgson J.B., Engel J.D.;  
 RT "The nucleotide sequence of the adult chicken alpha-globin genes.";  
 RL J. Biol. Chem. 258:4623-4629(1983).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82082384; PubMed=6273837;  
 RA Dodgson J.B., McCune K.C., Rusling D.J., Krust A., Engel J.D.;  
 RT "Adult chicken alpha-globin genes alpha A and alpha D: no anemic  
 RT shock alpha-globin exists in domestic chickens.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 78:5998-6002(1981).  
 RN [4]  
 RP PRELIMINARY SEQUENCE.  
 RX MEDLINE=76189993; PubMed=1225908;  
 RA Takei H., Ota Y., Wu K.C., Kiyohara T., Matsuda G.;  
 RT "Amino acid sequence of the alpha chain of chicken AI hemoglobin.";  
 RL J. Biochem. 77:1345-1347(1975).  
 RN [5]  
 RP SEQUENCE OF 1-63 AND 94-122.  
 RC TISSUE=Embryo;  
 RX MEDLINE=82098109; PubMed=7054172;  
 RA Chapman B.S., Hood L.E., Tobin A.J.;  
 RT "Minor early embryonic chick hemoglobin M. Amino acid sequences of  
 RT the epsilon and alpha D chains.";  
 RL J. Biol. Chem. 257:651-658(1982).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=99156923; PubMed=10037733;  
 RA Knapp J.E., Oliveira M.A., Xie Q., Ernst S.R., Riggs A.P.,  
 RA Hackert M.B.;  
 RT "The structural and functional analysis of the hemoglobin D component  
 RT from chicken.";  
 RL J. Biol. Chem. 274:6411-6420(1999).  
 CC -1- FUNCTION: Involved in oxygen transport from the lung to the  
 CC various peripheral tissues.  
 CC -1- SUBUNIT: Heterotetramer of two alpha-D chains and two beta chains.  
 CC The component D forms dimers of tetramers upon deoxygenation.  
 CC -1- TISSUE SPECIFICITY: Red blood cells.  
 CC -1- DEVELOPMENTAL STAGE: In birds, the alpha-D chain occurs in a minor  
 CC hemoglobin component, called hemoglobin d, which is expressed in  
 CC late embryonic and adult life.  
 CC -1- MISCELLANEOUS: REF.5 CHAIN WAS ISOLATED FROM HEM, THE LEAST  
 CC ABUNDANT OF THE FOUR EARLY CHICK HEMOGLOBINS.  
 CC -1- SIMILARITY: Belongs to the globin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC ENBL; X59989; CAA42605.1; --  
 CC ENBL; V00411; CAA23702.1; --  
 CC ENBL; M15378; AAA48584.1; ALT\_SEQ.  
 CC ENBL; J00853; AAA48800.1; --  
 CC PIR; B92421; EACHI.  
 CC PDS; IHR; 26-MAR-99.  
 CC InterPro; IPR002338; Alpha.haem.  
 CC InterPro; IPR000971; Globin.  
 CC Pfam; PF00042; Globin; 1.  
 CC PRINTS; PR00612; ALPHAHAE.  
 CC PROSITE; PS01033; GLOBIN; 1.  
 CC Heme; Oxygen transport; Transport; Erythrocyte;  
 CC 3D-structure.  
 CC METAL 58 58 IRON (HEME DISTAL LIGAND).  
 CC METAL 87 87 IRON (HEME PROXIMAL LIGAND).  
 CC CONFLICT 16 16 K -> R (IN REF. 1).  
 CC CONFLICT 107 107 V -> C (IN REF. 3 AND 5).  
 CC HELIX 4 17  
 CC TURN 18 19



```

FT HELIX 21 35
FT HELIX 37 42
FT TURN 44 45
FT TURN 50 51
FT HELIX 53 71
FT TURN 72 74
FT HELIX 76 79
FT TURN 81 88
FT HELIX 89 91
FT TURN 95 95
FT TURN 96 113
FT HELIX 114 116
FT HELIX 119 136
FT TURN 137 137
SQ SEQUENCE 141 AA; 15695 MW; 1FE426969B7B5384 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 625 GAGNAV 631
Db 64 GAGNAV 70

RESULT 54
HEAD MELGA STANDARD; PRT; 141 AA.
AC P81024;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hemoglobin alpha-D chain.
GN HEAD.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE.
RX MEDLINE=96066279; PubMed=7576241;
RA Eguchi Y., Ikehara T., Kayo S., Eguchi T., Takei H.;
RT "Amino acid sequence of alpha- and beta-polypeptide chains of turkey
(Meleagris gallopavo) hemoglobin.";
RL Biol. Chem. Hoppe-Seyler 376:437-440(1995).
CC -1- FUNCTION: Involved in oxygen transport from the lung to the
various peripheral tissues.
CC -1- SUBUNIT: Heterotetramer of two alpha-D chains and two beta chains.
CC -1- TISSUE SPECIFICITY: Red blood cells.
CC -1- DEVELOPMENTAL STAGE: In birds, the alpha-D chain occurs in a minor
hemoglobin component, called hemoglobin d, which is expressed in
late embryonic and adult life.
CC -1- SIMILARITY: Belongs to the globin family.
PIR; S56103; S56103.
DR HSSP; P02001; IHRP.
DR InterPro; IPR002338; Alpha_haem.
DR Pfam; PF000971; Globin.
DR Pfam; PF00042; globin; 1.
DR PRINTS; PR00612; ALPHAHAE.
DR PROSITE; PS01033; GLOBIN; 1.
KW Heme; Oxygen transport; Transport; Erythrocyte.
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA; 15665 MW; 59942696887954E7 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 625 GAGNAV 631
Db 64 GAGNAV 70

```

```

RESULT 55
MK_CHICK STANDARD; PRT; 142 AA.
AC P27052;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Mdkine precursor (Retinoic acid-induced heparin-binding protein)
DE (RI-HB).
GN RIHB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 80-85; 112-118 AND 121-133.
RX TISSUE=Embryo;
RX MEDLINE=91207359; PubMed=2018506;
RA Urios P., Duprez D., le Caer J.-P., Courtois Y., Vigny M., Laurent M.;
RT "Molecular cloning of RI-HB, a heparin binding protein regulated by
retinoic acid.";
RL Biochem. Biophys. Res. Commun. 175:617-624(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95010085; PubMed=7925417;
RA Duprez D., Treaggar J., Pecqueur C., Vigny M.R.;
RT "Organisation and promoter activity of the
retinoic-acid-induced-heparin-binding (RIHB) gene.";
RL Eur. J. Biochem. 224:931-941(1994).
RN [3]
RP SEQUENCE OF 22-77.
RX MEDLINE=91128406; PubMed=1993066;
RA Raulais D., Lagente-Chevallier O., Guettet C., Duprez D., Courtois Y.,
RA Vigny M.;
RT "A new heparin binding protein regulated by retinoic acid from chick
embryo.";
RL Biochem. Biophys. Res. Commun. 174:708-715(1991).
RN [4]
RP SEQUENCE OF 22-77.
RX STRAIN=White leghorn; TISSUE=Embryo;
RX MEDLINE=90108010; PubMed=2558016;
RA Vigny M., Raulais D., Puzenat N., Duprez D., Hartman M.P.,
RA Jeanny J.C., Courtois Y.;
RT "Identification of a new heparin-binding protein localized within
chick basement membranes.";
RL Eur. J. Biochem. 186:733-740(1989).
CC -1- FUNCTION: Has mitogenic activity, and neurite extension activity
for PC12 cells.
CC -1- SUBCELLULAR LOCATION: BASEMENT MEMBRANES IN EARLY EMBRYONIC
TISSUES, AND CELL SURFACE OF NEUROECTODERMAL CELLS.
CC -1- DEVELOPMENTAL STAGE: Essentially expressed during embryogenesis.
CC -1- INDUCTION: By retinoic acid.
CC -1- SIMILARITY: Belongs to the pleiotrophin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC DR EMBL; M61754; -; NOT ANNOTATED_CDS.
CC DR EMBL; X76482; CAA54020.1; -.
CC PIR; J00573; J00573.
CC HSSP; P21741; LMKC.
CC InterPro; IPR000762; PTN_MK.
CC Pfam; PF01091; PTN_MK_C; 1.
CC Pfam; PF05196; PTN_MK_N; 1.
CC PRINTS; PR00269; PTNIDKINE.
CC ProDom; PD005592; PTN_MK; 1.

```

DR SMART: SM00193; PTN: 1;  
DR PROSITE; PS00619; PTN MK\_1; 1;  
DR PROSITE; PS00620; PTN MK\_2; 1;  
KW Growth factor; Mitogen; Differentiation; Heparin-binding; Signal.  
FT SIGNAL 1 21  
FT CHAIN 22 142 MIDKINE.  
FT DISULFID 36 60 BY SIMILARITY.  
FT DISULFID 44 69 BY SIMILARITY.  
FT DISULFID 51 73 BY SIMILARITY.  
FT DISULFID 83 115 BY SIMILARITY.  
FT DISULFID 93 125 BY SIMILARITY.  
FT CONFLICT 89 89 S -> R (IN REF. 2).  
FT CONFLICT 93 93 C -> G (IN REF. 2).  
SQ SEQUENCE 142 AA; 15579 MW; 9D05CAFB9558451B CRC64;  
  
Query Match 0.68; Score 7; DB 1; Length 142;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 7; Conservative 0; Mismatches 0; Caps 0;  
  
Qy 1153 LLLALL 1159  
Db 7 LLLALL 13  
  
RESULT 56  
BIK\_HUMAN STANDARD; PRT; 150 AA.  
ID BIK\_HUMAN  
AC Q13523; Q16582;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Bcl-2 interacting killer (Apoptosis inducer NBK) (Bp4) (BIP1).  
GN BIK OR NBK.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutaria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=B-cell;  
RX MEDLINE=96068922; PubMed=7478623;  
RA Boyd J.M., Gallo G.J., Elangovan B., Houghton A.B., Malstrom S.,  
RA Avery B.J., Ebb R.G., Subramanian T., Chittenden T., Lutz R.J.,  
RA Chinnadurai G.;  
RT "Bik", a novel death-inducing protein shares a distinct sequence motif  
RT with Bcl-2 family proteins and interacts with viral and cellular  
RT survival-promoting proteins.  
RL Oncogene 11:1921-1928(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96413338; PubMed=8816500;  
RA Han J., Sabbatini P., White E.;  
RT "Induction of apoptosis by human NBK/Bik, a BH3-containing protein  
RT that interacts with E1B 19k";  
RL Mol. Cell. Biol. 16:5857-5864(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymphoid;  
RA Pun K.-T., Farrow S.N., Raven T., Wride C.J., White J.H.M., Brown R.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99431905; PubMed=10500065;  
RA Castellani A., Ino Y., Louis D.N., Ramesh V., Gusella J.F., Rustgi A.K.;  
RT "Mapping of a target region of allelic loss to a 0.5-cM interval on  
RT chromosome 22q13 in human colorectal cancer."  
RL Gastroenterology 117:831-837(1999).  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Rieder M.J., Livingston R.J., Daniels M.R., Chung M.-W.,  
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,  
RA Schackwitz W.S., Sherwood J.K., Wittrak L.A., Nickerson D.A.;  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.

[6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20057165; PubMed=10591208;  
RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,  
RA Clump M., Smink L.J., Ainscough R., Almeida J.P., Sabaghe A.K.,  
RA Baguley C., Bailey J., Barlow K.P., Bates K.N., Beasley O.P.,  
RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,  
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,  
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,  
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson B.,  
RA Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,  
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,  
RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,  
RA Hunt R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,  
RA Hunt S.B., Jones M.C., Kershaw J., Kimberley A.M., King A.,  
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,  
RA Marjyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,  
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,  
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,  
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,  
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,  
RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,  
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,  
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,  
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudon J.,  
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,  
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,  
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,  
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,  
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,  
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,  
RA Zhan M., Zhang G., Chisoe S., Murray J.J., Miller N., Minx P.,  
RA Fulton R., Johnson D., Hemis G., Bentley D., Bradshaw H., Bourne S.,  
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,  
RA Hinds K., Kemp K., Latreille P., Layman D., Ozerky P., Rohlfing T.,  
RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,  
RA Korf I., Bedell J.A., Hillier L.M., Mardis E., Waterston R.,  
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,  
RA Sudar M.I., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,  
RA Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,  
RA Peyzard M., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,  
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,  
RA Khan A.S., Lane L., Tilahun Y., Wright H.;  
RT "The DNA sequence of human chromosome 22."  
RL Nature 402:489-495(1999).  
RN [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymph;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Rutherford Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [8]  
RP MUTAGENESIS, AND FUNCTION OF BH3 DOMAIN.  
RX MUTAGENESIS, AND FUNCTION OF BH3 DOMAIN.  
RN [8]  
RP MUTAGENESIS, AND FUNCTION OF BH3 DOMAIN.  
RX MUTAGENESIS, AND FUNCTION OF BH3 DOMAIN.

```

RA Chittenden T., Flemington C., Houghton A.B., Ebb R.G., Gallo G.J.,
RA Elangovan B., Chinnadurai G., Lutz R.J.;
RT "A conserved domain in Bak, distinct from BH1 and BH2, mediates cell
RT death and protein binding functions.";
RL EMBO J. 14:5589-5596(1995)
CC -!- FUNCTION: Accelerates programmed cell death. Binding to the
CC apoptosis repressors Bcl-X(L), BHRF1, Bcl-2 or its adenovirus
CC homolog E1B 19k protein suppresses this death-promoting activity.
CC Does not interact with BAX.
CC -!- SUBCELLULAR LOCATION: Around the nuclear envelope, and in
CC cytoplasmic membranes.
CC -!- DOMAIN: Interact BH3 domain is required by BIK, BID, BAX, BAD and
CC BAX for their pro-apoptotic activity and for their interaction
CC with anti-apoptotic members of the Bcl-2 family.
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U34584; AAC50413.1; -
DR EMBL; U49730; AAC79124.1; -
DR EMBL; X89986; CA862013.1; -
DR EMBL; AF174424; AAF01156.1; -
DR EMBL; AF174421; AAF01156.1; JOINED.
DR EMBL; AF174422; AAF01156.1; JOINED.
DR EMBL; AF174423; AAF01156.1; JOINED.
DR EMBL; AY245248; AAO61089.1; -
DR EMBL; AL022237; CAA18260.2; -
DR EMBL; BC001599; RAH01599.1; -
DR EMBL; S89214; S58214.
DR EMBL; HGNC:1051; BIK.
DR MIM; 603392; -
DR GO; GO:0008632; P:apoptotic program; TAS.
DR GO; GO:0006917; P:induction of apoptosis; TAS.
DR InterPro; IPR000712; Bcl2_BH.
DR PROSITE; PS01259; BH3; 1.
KW Apoptosis; Transmembrane.
FT DOMAIN 57 71 BH3.
FT TRANSMEM 136 156 POTENTIAL.
FT DOMAIN 137 158 LEUCINE-ZIPPER (POTENTIAL).
FT CONFLICT 149 150 PL -> LP (IN REF. 1).
SQ SEQUENCE 160 AA; 18016 MW; 89034F443F5A136 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLL 1159
DB 141 LLLLLL 147

RESULT 57
NUOE AQUAE
ID NUOE AQUAE STANDARD; PRT; 160 AA.
AC O66842.
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NADH-quinone oxidoreductase chain E (EC 1.6.99.5) (NADH dehydrogenase
DE I, chain E) (NDH-1, chain E).
GN NUOE OR AQ 574.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;

```

```

RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.B., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -!- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
CC sulfur (Fe-S) centers, to quinones in the respiratory chain.
CC Couples the redox reaction to proton translocation (for every two
CC electrons transferred, four hydrogen ions are translocated across
CC the cytoplasmic membrane), and thus conserves the redox energy in
CC a proton gradient (by similarity).
CC -!- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -!- COFACTOR: Binds 1 2Fe-2S cluster (Potential).
CC -!- SIMILARITY: Belongs to the complex I 24 kDa subunit family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000696; AAC06799.1; -
DR PIR; P70351; P70351.
DR InterPro; IPR002023; Cmplx1_24kDa.
DR Pfam; PF01257; complex1_24kDa; 1.
DR ProDom; PD003859; Cmplx1_24kDa; 1.
DR PROSITE; PS01099; COMPLEX1_24K; 1.
KW Oxidoreductase; NAD; Quinone; Metal-binding; Iron-sulfur; Iron;
FT METAL 86 86 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 91 91 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 127 127 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 131 131 IRON-SULFUR (2FE-2S) (POTENTIAL).
SQ SEQUENCE 160 AA; 18550 MW; 425D81995A491B3B CRC64;

Query Match 0.6%; Score 7; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 413 EFPEELK 419
DB 7 EFPEELK 13

RESULT 58
TCH2 ARATH
ID TCH2 ARATH STANDARD; PRT; 161 AA.
AC P25070; O22592;
DT 01-MAY-1992 (Rel. 22, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Calmodulin-related protein 2, touch-induced.
GN TCH2 OR AR5G3770 OR K22F20.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Johnson K.A., Braam J.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=99087489; PubMed=9872454;
RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
RA Tabata S.;

```

RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.  
 RT Sequence features of the regions of 1,013,767 bp covered by sixteen  
 RT physically assigned P1 and TAC clones."  
 RL DNA Res. 5:297-308(1998).  
 RN [3]  
 RP SEQUENCE OF 26-70 FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=90150263; PubMed=2302732;  
 RA Braam J., Davis R.W.;  
 RT "Rain-, wind-, and touch-induced expression of calmodulin and  
 RT calmodulin-related genes in Arabidopsis."  
 RL Cell 60:357-364(1990).  
 RN [4]  
 RP 3D-STRUCTURE MODELING OF 7-158.  
 RX MEDLINE=97189489; PubMed=9037719;  
 RA Zhan A.R.; Johnson K.A.; Braam J., James M.N.G.;  
 RT "Comparative modeling of the three-dimensional structure of the  
 RT calmodulin-related TCH2 protein from Arabidopsis."  
 RL Proteins 27:144-153(1997).  
 CC -!- INDUCTION: By rain-, wind-, and touch (thigmomorphogenesis).  
 CC -!- SIMILARITY: Contains 4 EF-hand calcium-binding domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AB026473; AB82713.1; -;  
 DR EMBL; AB016873; BAB10353.1; -;  
 DR PIR; C34669; C34669.  
 DR PDB; 1AVJ; 28-JAN-98.  
 DR InterPro; IPR003299; Calflagin.  
 DR InterPro; IPR002048; EF-hand.  
 DR Pfam; PF00036; ehand; 4  
 DR PRINTS; PR01362; CALFLAGIN.  
 DR PRODOM; PD000012; EF-hand; 2.  
 DR SMART; SM00054; EFh; 4.  
 DR PROSITE; PS00018; EF HAND; 4.  
 DR Calcium-binding; Repeat; 3D-structure.  
 KW CA\_BIND 26 37 EF-HAND 1 (POTENTIAL).  
 FT CA\_BIND 62 73 EF-HAND 2 (POTENTIAL).  
 FT CA\_BIND 103 114 EF-HAND 3 (POTENTIAL).  
 FT CA\_BIND 139 150 EF-HAND 4 (POTENTIAL).  
 FT CONFLICT 54 54 T -> Y (IN REF. 3).  
 FT HELIX 12 25  
 FT TURN 27 28  
 FT STRAND 33 33  
 FT HELIX 35 44  
 FT TURN 45 45  
 FT HELIX 51 61  
 FT TURN 63 64  
 FT STRAND 69 69  
 FT HELIX 71 102  
 FT TURN 104 105  
 FT STRAND 110 110  
 FT HELIX 112 121  
 FT TURN 122 123  
 FT HELIX 128 138  
 FT STRAND 146 146  
 FT HELIX 148 156  
 SQ SEQUENCE 161 AA; 17546 MW; A7324A5C0B3CDB9 CRC64;  
 Query Match 0.6%; Score 7; DB 1; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 87;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 487 VDIDGDG 493  
 DB 138 VDIDGDG 144

RESULT 59  
 YF87 METJA STANDARD; PRT; 171 AA.  
 AC Q58982;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein MJ1587.  
 GN MJ1587.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 OC Methanocaldococcales; Methanocaldococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=868087;  
 RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kervatage A.R., Dougherty B.A., Tomb J.-P., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.D., Sadow P.W., Borodovsky M.,  
 RA Usterback T.R., Kelley J.M., Peterson J.D., Hurst M.A., Kaine B.P.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jannaschii."  
 RL Science 273:1058-1073(1996).  
 CC -!- SIMILARITY: TO M.JANNASCHII MJ0417.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U67599; AAB99616.1; -;  
 DR PIR; B64498; B64498.  
 DR TIGR; MJ1587;  
 DR InterPro; IPR003141; PHP N.  
 DR Pfam; PF02231; PHP N; 1.  
 DR SMART; SM00481; POLIILAC; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 171 AA; 19198 MW; 7D4F7B936C5C7472 CRC64;  
 Query Match 0.6%; Score 7; DB 1; Length 171;  
 Best Local Similarity 100.0%; Pred. No. 91;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 344 ALGDRIF 350  
 DB 123 ALGDRIF 129  
 RESULT 60  
 YD2F SCHPO STANDARD; PRT; 176 AA.  
 AC Q10263;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Very hypothetical protein C56F8.15 in chromosome I.  
 GN SPAC5F8.15.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.

```

RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feitwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle Z.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkmar G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Sucks M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borszym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z69728; CAA93586.1; -.
DR PIR; T38925; T38925.
DR GeneDB SPombe; SPAC56F8.15; -.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 27 47 POTENTIAL.
FT TRANSMEM 79 99 POTENTIAL.
FT TRANSMEM 107 127 POTENTIAL.
SQ SEQUENCE 176 AA; 20595 MW; B0FB48B199DA622A CRC64;

Query Match 0.6%; Score 7; DB 1; Length 176;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 820 YLSPDT 826
DB 138 YLSPDT 144

RESULT 61
HIS7 CHL7E
AC Q8KEF4; STANDARD; PRT; 194 AA.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19) (IGPD).
GN HISB OR CT0735.
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901;

Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.I., Yang P.,
Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
Nierman W.C., Feldblyum I.V., Hansen C.D., Craven M.B., Radune D.,
Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
Renteria J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT phototrophic, anaerobic, green-sulfur bacterium."
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
CC
CC -!- CATALYTIC ACTIVITY: D-erythro-1-(imidazol-4-yl)glycerol 3-
CC phosphate = 3-(imidazol-4-yl)-2-oxopropyl phosphate + H(2)O.
CC
CC -!- PATHWAY: Histidine biosynthesis; sixth step.
CC
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC
CC -!- SIMILARITY: Belongs to the imidazoleglycerol-phosphate dehydratase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB012844; AM71972.1; -.
DR TIGR; CT0735; -.
DR HAWAP; MF 00076; -.
DR InterPro; IPR000807; IGPD.
DR Pfam; PF00475; IGPD; 1.
DR ProDom; PD002282; IGPD; 1.
DR PROSITE; PS00954; IGP_DEHYDRATASE_1; 1.
DR PROSITE; PS00955; IGP_DEHYDRATASE_2; 1.
KW Histidine biosynthesis; lyase; Complete proteome.
SQ SEQUENCE 194 AA; 21189 MW; D87295AAC9B830F2 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 194;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 IVDALGD 347
DB 80 IVDALGD 86

RESULT 62
Z239 MOUSE
ID Z239_MOUSE STANDARD; PRT; 201 AA.
AC P24399;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 239 (Zfp-239) (Zinc finger protein MOK-2).
GN ZNF239 OR ZFP239 OR MOK2 OR MOK-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90097859; PubMed=2104662;
RA Ennault-Lange M., Kress M., Hamer D.;
RT "A gene that encodes a protein consisting solely of zinc finger
RT domains is preferentially expressed in transformed mouse cells."
RL Mol. Cell. Biol. 10:418-421(1990).
CC
CC -!- FUNCTION: May function as a transcription factor.
CC
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC
CC -!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN TRANSFORMED
CC MOUSE CELLS.
CC
CC -!- SIMILARITY: BELONGS TO THE KRUPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC
CC -!- SIMILARITY: Contains 7 C2H2-type zinc fingers.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; M32057; AAA39729.1; -;  
 DR PIR; I57505; I57505.  
 DR HSSP; P08047; I5P2.  
 DR TRANSFAC; T00510; -;  
 DR MGD; MGI:1306812; Zfp239.  
 DR InterPro; IPR007087; Znf\_C2H2.  
 DR InterPro; IPR007086; Znf\_C2H2\_sub.  
 DR Pfam; PF00096; Zf-C2H2; 7.  
 DR PRINTS; PR00048; ZINC\_FINGER.  
 DR PRODOM; PD000003; Znf\_C2H2; 6.  
 DR SMART; SM00355; Znf\_C2H2; 7.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 7.  
 DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 7.  
 KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;  
 KW Nuclear protein; Repeat.  
 FT ZN\_FING 6 28 C2H2-TYPE 1.  
 FT ZN\_FING 34 56 C2H2-TYPE 2.  
 FT ZN\_FING 62 84 C2H2-TYPE 3.  
 FT ZN\_FING 90 112 C2H2-TYPE 4.  
 FT ZN\_FING 118 140 C2H2-TYPE 5.  
 FT ZN\_FING 146 168 C2H2-TYPE 6.  
 FT ZN\_FING 174 196 C2H2-TYPE 7.  
 SQ SEQUENCE 201 AA; 22832 MW; 57A18C0AADA1384 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 959 FTRSSSL 965  
 DB 15 FTRSSSL 21

## RESULT 63

ID GPBB\_MOUSE STANDARD; PRT; 206 AA.  
 AC P56400;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Platelet glycoprotein Ib beta chain precursor (GP-Ib beta) (GP1BB)  
 DE (GP1b-beta).  
 GN GP1BB.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97403789; PubMed=9259114;  
 RA Kitaguchi T., Murata M., Anbo H., Moriki T., Ikeda Y.;  
 RT "Characterization of the gene encoding mouse platelet glycoprotein Ib  
 RT beta.";  
 RL Thromb. Res. 87:235-244(1997).  
 CC -!- FUNCTION: GP-Ib, a surface membrane protein of platelets,  
 CC participates in the formation of platelet plugs by binding to von  
 CC Willebrand factor, which is already bound to the subendothelium  
 CC (by similarity).  
 CC -!- SUBUNIT: GP-Ib alpha and beta are disulfide linked. GP-IX is  
 CC complexed with the GP-Ib heterodimer via a non covalent linkage  
 CC (by similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- MISCELLANEOUS: Platelet activation apparently involves disruption  
 CC of the macromolecular complex of GP-Ib with the platelet  
 CC glycoprotein IX (GP-IX) and dissociation of GP-Ib from the actin-

CC binding protein.  
 CC -!- SIMILARITY: Contains 1 leucine-rich (LRR) repeat.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; AB001419; BAA22424.1; -;  
 DR MGD; MGI:107852; Gp1bb.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000483; LRR\_Cterm.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR Pfam; PF00560; LRR; 1.  
 DR Pfam; PF01463; LRRCT; 1.  
 DR Pfam; PF01462; LRRNT; 1.  
 DR SMART; SM00082; LRRCT; 1.  
 DR SMART; SM00013; LRRNT; 1.  
 KW Platelet; Transmembrane; Glycoprotein; Hemostasis; Blood coagulation;  
 KW Signal; Phosphorylation; Cell adhesion; Leucine-rich repeat.  
 FT SIGNAL 1 26 BY SIMILARITY.  
 FT CHAIN 27 206 PLATELET GLYCOPROTEIN IB BETA CHAIN.  
 FT DOMAIN 27 147 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 148 172 POTENTIAL.  
 FT DOMAIN 173 206 CYTOPLASMIC (POTENTIAL).  
 FT REPEAT 60 83 LRR.  
 FT CARBOHYD 66 66 N-LINKED (GLCNAc... ) (POTENTIAL).  
 FT MOD RES 191 191 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).  
 SQ SEQUENCE 206 AA; 21762 MW; AC4BCB4DFA226FID CRC64;

Query Match 0.6%; Score 7; DB 1; Length 206;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALL 1159  
 DB 13 LLLALL 19

## RESULT 64

ID PSMB\_SULSO STANDARD; PRT; 208 AA.  
 AC Q9UXF3;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Proteasome beta subunit precursor (EC 3.4.25.1) (Multicatalytic  
 DE endopeptidase complex beta subunit).  
 GN PSMB OR SSO0766 OR C40\_002.  
 OS Sulfolobus solfataricus.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OC Sulfolobus.  
 OX NCBI\_TaxID=2287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ATCC 35092 / DSM 1617 / P2;  
 RX MEDLINE=20165948; PubMed=10701121;  
 RA Charlebois R.L., Singh R.K., Chan-Weher C.C.-Y., Allard G., Chow C.,  
 RA Confalonieri T., Curtis B., Duguet M., Brauso G., Faguy D.,  
 RA Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,  
 RA Kuehwa N., Lafleur E., Medina N., Peng X., Penny S.L., She Q.,  
 RA St Jean A., van der Oost J., Young P., Zivanovic Y., Doolittle W.F.,  
 RA Ragan M.A., Sengen C.W.;  
 RT "Gene content and organization of a 281-kbp contig from the genome of  
 RT the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";  
 RL Genome 43:116-136(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ATCC 35092 / DSM 1617 / P2;  
 RX MEDLINE=21332296; PubMed=11427726;

```

RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aways M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Brauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;
RA "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -!- FUNCTION: The proteasome is a multicatalytic proteinase complex
CC which is characterized by its ability to cleave peptides with Arg,
CC Phe, Tyr, Leu, and Glu adjacent to the leaving group at neutral or
CC slightly basic pH (By similarity).
CC -!- CATALYTIC ACTIVITY: Cleavage at peptide bonds with very broad
CC specificity.
CC -!- SUBUNIT: Composed of two subunits, alpha and beta. The complex is
CC formed of four rings. The two outer rings are each composed of
CC seven alpha subunits. The two inner rings are each composed of
CC seven beta subunits (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to peptidase family T1B.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y18930; CAB57537.1; ALT INIT.
CC EMBL; AE006700; AAK41063.1; ALT_INIT.
CC HSSP; P28061; LPMA.
CC MEROPS; T01.002; -.
CC InterPro; IPR000243; Pept_T1A_subb.
CC InterPro; IPR001353; Peptidase_T1.
CC Pfam; PF00227; Proteasome; 1.
CC PRINTS; PR00141; PROTEASOME.
CC PROSITE; PS00854; PROTEASOME B; 1.
CC Proteasome; Hydrolase; Protease; Complete proteome;
KW Threonine protease.
FT PROPEP 1 14 APPARENTLY REMOVED IN MATURE FORM (BY
FT SIMILARITY).
FT CHAIN 15 208 PROTEASOME BETA SUBUNIT.
FT ACT SITE 15 15 PROBABLE.
SQ SEQUENCE 208 AA; 22927 MW; C940A8273AB25CB9 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 685 TTTVGIR 692
DB 15 TTTVGIR 21

RESULT 65
KPTA_PYRAE STANDARD; PRT; 213 AA.
AC Q8ZSP2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable RNA 2'-phosphotransferase (EC 2.7.7.-).
GN KPTA OR PAE3647.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;

RA Fitz-Gibbon S.F., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RA "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RA aerophilum."
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
CC -!- FUNCTION: Removes the 2'-phosphate from RNA via an intermediate in
CC which the phosphate is ADP-ribosylated by NAD followed by a
CC presumed transesterification to release the RNA and generate ADP-
CC ribose 1',2'-cyclic phosphate (APPR>P). May function as an ADP-
CC ribosylase (By similarity).
CC -!- SIMILARITY: Belongs to the kptA / TPT1 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE009943; AAL65071.1; -.
CC HAMAP; MF_00299; -.
CC InterPro; IPR002745; Ptrans_KptA/Tpt1.
CC Pfam; PF01885; PTS_2-RNA; 1.
CC ProDom; PD009619; P04transfrs; 1.
CC Transfrase; NAD; Complete proteome.
SQ SEQUENCE 213 AA; 23452 MW; 12DF2AD12CE69636 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 DVYKCPV 78
DB 3 DVYKCPV 9

RESULT 66
CH13_HUMAN STANDARD; PRT; 214 AA.
AC Q9SKS9; Q9N3M3; Q9NSR0.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C8orf13.
GN C8ORF13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Appel S., Bergheim A., Reichwald K., Reis A., Rosenthal A., Ramsay M.,
RA Hennies H.;
RA "Transcript map of the KWE critical region on chromosome 8p22-p23."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Amalgam;
RA Wambutt R., Heubner D., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 83-214 FROM N.A.
RC TISSUE=Amalgam;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SRC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ301564; CAC82740.1; -  
DR EMBL; AL834122; CAD38843.1; -  
DR EMBL; AL157475; CAB75670.1; -  
DR PIR; T46905; T46905.  
DR Genew; HGNC:15549; Csoxf13.  
KW Hypothetical protein; Coiled coil.  
FT DOMAIN 118 156 COILED COIL (POTENTIAL).  
FT CONFLICT 56 56 H -> Q (IN REF. 2).  
FT CONFLICT 107 107 T -> S (IN REF. 3).  
FT CONFLICT 130 130 M -> L (IN REF. 2).  
SQ SEQUENCE 214 AA; 24182 MW; 14216458987284 CRC64;  
  
Query Match 0.6%; Score 7; DB 1; Length 214;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1090 LRSUKAL 1096  
DB 27 LRSUKAL 33  
  
RESULT 67  
CYB CERCE  
ID CYB CERCE STANDARD; PRT; 214 AA.  
AC P87419;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cytochrome b (Fragment).  
GN MTCYB OR COB OR CYTB.  
OS Cerastes cerastes (Horned desert viper).  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;  
OC Viperidae; Viperinae; Cerastes.  
OX NCBI\_TaxID=8697;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98334559; PubMed=9667984;  
RA Vidal N., Lecoindre G.;  
RT "Weighting and congruence: a case study based on three mitochondrial genes in pitvipers."  
RL Mol. Phylogenet. Evol. 9:366-374(1998).  
RN [2]  
  
RP SEQUENCE OF 1-132 FROM N.A.  
RA Vidal N., Lecoindre G., Vie J.-C., Gasc J.-P.;  
RT "Molecular systematics of pitvipers: paraphyly of the Bothrops complex."  
RL C. R. Acad. Sci., III, Sci. Vie 320:95-101(1997).  
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis (By similarity).  
CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or b562) is low-potential and absorbs at about 562, and heme 2 (or BH or b566) is high-potential and absorbs at about 566 (By similarity).  
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b, cytochrome c1 and the Rieske protein (By similarity).  
CC -!- SIMILARITY: Belongs to the cytochrome b family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF039265; AAC33542.1; -

DR InterPro; IPR005798; Cytb\_b6\_C.  
DR InterPro; IPR005797; Cytb\_b6\_N.  
DR Pfam; PF00033; Cytochrome b\_N; 1.  
DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; FALSE NEG.  
DR PROSITE; PS00193; CYTOCHROME\_B\_CO; PARTIAL.  
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;  
KW Heme.  
FT NON\_TER 1 1  
FT METAL 81 81 IRON 1 (HEME B562 AXIAL LIGAND).  
FT METAL 95 95 IRON 2 (HEME B566 AXIAL LIGAND).  
FT NON\_TER 214 214  
SQ SEQUENCE 214 AA; 24230 MW; 755818DB9204P820 CRC64;  
  
Query Match 0.6%; Score 7; DB 1; Length 214;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 37 RTAFPGY 43  
DB 123 RTAFPGY 129  
  
RESULT 68  
FGF3 CHICK  
ID FGF3 CHICK STANDARD; PRT; 220 AA.  
AC P48801;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fibroblast growth factor-3 precursor (FGF-3) (HBGF-3).  
GN FGF3 OR FGF-3.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Rhode Island red; TISSUE=Embryo;  
RC MEDLINE=95309122; PubMed=7789270;  
RA Mahmood R., Kiefer P., Guthrie S., Dickson C., Mason I.;  
RT "Multiple roles for FGF-3 during cranial neural development in the chicken."  
RL Development 121:1399-1410(1995).  
CC -!- FUNCTION: POTENT MITOGEN AND TRANSFORMING AGENT (BY SIMILARITY).  
CC -!- SIMILARITY: Belongs to the heparin-binding growth factors family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z47555; CAA87635.1; -  
DR PIR; I50588; I50588.  
DR HSP; P31371; IG82.  
DR InterPro; IPR008996; Cytok\_IL1\_like.  
DR InterPro; IPR002348; IL1\_HBGF.  
DR Pfam; PF00167; FGF; 1.  
DR PRINTS; PR00262; IL1HBGF.  
DR ProDom; PD000831; IL1\_HBGF; 1.  
DR SMART; SM00442; FGF; 1.  
DR PROSITE; PS00247; HBGF\_FGF; 1.  
KW Growth factor; Mitogen; Signal; Glycoprotein.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 220 FIBROBLAST GROWTH FACTOR-3.  
FT CARBOHYD 66 66 N-LINKED (GLCNAC..) (POTENTIAL).  
SQ SEQUENCE 220 AA; 25050 MW; B15D41D1E551C5D5 CRC64;  
  
Query Match 0.6%; Score 7; DB 1; Length 220;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;



Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALL 1159  
 Db 6 LLLALL 12

RESULT 69  
 SDFL HUMAN  
 ID SDFL HUMAN STANDARD; PRT; 221 AA.  
 AC Q9HCN8; Q9BR15;  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Stromal cell-derived factor 2-like protein 1 precursor (SDF2 like  
 DE protein 1) (PwP1-interacting protein 8).  
 GN SDF2L1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=21092619; PubMed=11162531;  
 RA Fukuda S., Sumii M., Masuda Y., Takahashi M., Koike N., Teishima J.,  
 RA Yasumoto H., Itamoto T., Asahara T., Dohi K., Kamiya K.;  
 RT "Murine and human SDF2L1 is an endoplasmic reticulum stress-inducible  
 RT gene and encodes a new member of the Pmt/rt protein family.";  
 RL Biochem. Biophys. Res. Commun. 280:407-414(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RA Honore B.;  
 RT "PwP1-interacting protein 8";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=2388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshikiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.W., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
 RA Schnerch A., Schein J.E., Jones S.D.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (Potential).  
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed with high expression in  
 CC testis, moderate expression in the pancreas, spleen, prostate,  
 CC small intestine and colon. Very low expression is seen in brain  
 CC and skeletal muscle.  
 CC -1- SIMILARITY: Contains 3 MIR domains.

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC CC  
 DR EMBL; AB043007; BAB18277.1; -  
 DR EMBL; AF277316; AAK69113.1; -  
 DR EMBL; BC006248; AAH06248.1; -  
 DR PIR; JC7587; JC7587.  
 DR Genew; HGNC:10676; SDF2L1.  
 DR MIN; 607551; -  
 DR InterPro; IPR000886; ER\_target\_S.  
 DR InterPro; IPR003608; MIR.  
 DR Pfam; PF02815; MIR; 3.  
 DR SMART; SMO0472; MIR; 3.  
 DR PROSITE; PS00014; ER\_TARGET; 1.  
 DR PROSITE; PS00919; MIR; 3.  
 DR Endoplasmic reticulum; Signal; Repeat.  
 KW POTENTIAL.  
 FT SIGNAL 1 28  
 FT CHAIN 29 221  
 FT STROMAL CELL-DERIVED FACTOR 2-LIKE  
 FT PROTEIN 1.  
 FT DOMAIN 33 87  
 FT MIR 1.  
 FT DOMAIN 95 150  
 FT MIR 2.  
 FT DOMAIN 151 205  
 FT MIR 3.  
 FT SITE 218 221  
 FT PREVENT SECRETION FROM ER (POTENTIAL).  
 FT CONFLICT 94 94  
 FT C -> R (IN REF. 2 AND 3).  
 FT CONFLICT 162 162  
 FT L -> F (IN REF. 2 AND 3).  
 SQ SEQUENCE 221 AA; 23511 MW; 23EBEE14BE9E2EF CRC64;

Query Match 0.6% Score 7; DB 1; Length 221;  
 Best Local Similarity 100.0%; Pred.No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLV 1160  
 Db 17 LLLALLV 23

RESULT 70  
 EMBL\_MOUSE  
 ID EMBL\_MOUSE STANDARD; PRT; 223 AA.  
 AC Q61878;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Eosinophil granule major basic protein precursor (MBP) (Proteoglycan  
 DE 2, bone marrow).  
 GN PRG2 OR MBP-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN=CBA/J; TISSUE=Liver;  
 RX MEDLINE=95403975; PubMed=7673718;  
 RA Larson K.A., Horton M.A., Madden B.J., Gleich G.J., Lee N.A.,  
 RA Lee J.J.;  
 RT "The identification and cloning of a murine major basic protein gene  
 RT expressed in eosinophils";  
 RL J. Immunol. 155:3002-3012(1995).  
 CC -1- FUNCTION: Cytotoxin and histaminotoxin. MBP also induces  
 CC noncytolytic histamine release from basophils. It is involved in  
 CC antiparasitic defense mechanisms and immune hypersensitivity  
 CC reactions (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific  
 CC granule (crystalloid core) (By similarity).  
 CC -1- SIMILARITY: Contains 1 C-type lectin family domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

```

CC EMBL; L45768; AAA83027.1; --
CC HSP; P13727; I88U.
CC MGD; WGI.103294; PrG2.
CC InterPro; IPR002352; Emaior_basic.
CC InterPro; IPR001304; Lectin_C.
CC Pfam; PF00059; lectin_c; 1.
CC PRINTS; PR00770; EMAJORBASIC.
CC SMART; SM00034; CLECT; 1.
CC PROSITE; PS00615; C-TYPE LECTIN 1; FALSE_NEG.
CC PROSITE; PS00041; C-TYPE LECTIN 2; 1.
CC Eosinophil; Signal; Immune response; Antibiotic; Lectin.
FT SIGNAL 1 16
FT PROPEP 17 106
FT CHAIN 107 223
FT DOMAIN 124 223
FT DISULFID 126 221
FT DISULFID 198 213
FT DISULFID 223 AA; 24255 MW; 7D66D946DCEADA00 CRC64;
SQ SEQUENCE 223 AA; 24255 MW; 7D66D946DCEADA00 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLV 1160
Db 5 LLLALLV 11

RESULT 71
EMBP_RAT
ID EMBL_CAVPO STANDARD; PRT; 227 AA.
AC Q63189;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eosinophil granule major basic protein precursor (MBP).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Bone marrow;
RX MEDLINE=56138543; PubMed=8547309;
RA Nittoh T., Watanabe M., Okayama H., Misawa S., Isobe Y., Hayashi H.,
RA Mue S., Chuchi K.;
RT "Cloning of cDNA for rat eosinophil major basic protein.";
RL Biochim. Biophys. Acta 1264:261-264(1995).
CC -!- FUNCTION: Cytotoxic and helminthotoxin. MBP also induces
CC noncytolytic histamine release from basophils. It is involved in
CC antiparasitic defense mechanisms and immune hypersensitivity
CC reactions (By similarity).
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
CC granule (crystalloid core) (By similarity).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D50568; BAA09129.1; --
CC HSP; P13727; I88U.
CC InterPro; IPR002352; Emaior_basic.
CC InterPro; IPR001304; Lectin_C.
CC Pfam; PF00059; lectin_c; 1.
CC PRINTS; PR00770; EMAJORBASIC.

```

```

DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
DR PROSITE; PS00041; C-TYPE LECTIN 2; 1.
KW Eosinophil; Signal; Immune response; Antibiotic; Lectin.
FT SIGNAL 1 16
FT PROPEP 17 110
FT CHAIN 111 227
FT DOMAIN 128 227
FT DISULFID 130 225
FT DISULFID 202 217
FT DOMAIN 59 68
FT SEQUENCE 227 AA; 25129 MW; 34F2B8514090DE82 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 227;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLV 1160
Db 5 LLLALLV 11

RESULT 72
EMBL_CAVPO
ID EMBL_CAVPO STANDARD; PRT; 233 AA.
AC P22032;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eosinophil granule major basic protein 1 precursor (MBP-1).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 115-162.
RC TISSUE=Eosinophil;
RX MEDLINE=91160746; PubMed=1705901;
RA Aoki I., Shindoh Y., Nishida T., Nakai S., Hong Y.-M., Mio M.,
RA Saito T., Tasaka K.;
RT "Sequencing and cloning of the cDNA of guinea pig eosinophil major
RT basic protein.";
RL FEBS Lett. 279:330-334(1991).
RN [2]
RP PARTIAL SEQUENCE.
RX MEDLINE=94092714; PubMed=8268206;
RA Hashimoto Y., Nagaoka I., Yamashita T.;
RT "Purification of the antibacterial fragments of guinea-pig major
RT basic protein.";
RL Biochim. Biophys. Acta 1203:236-242(1993).
CC -!- FUNCTION: MBP may play some important roles in the allergic
CC reactions and inflammations, since MBP is capable of releasing
CC histamine from mast cells and damaging the epithelial cells of
CC bronchial tubes. Antiparasitic and antibiotic.
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
CC granule (crystalloid core).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90251; BAA14291.1; --
CC PIR; S13625; S13625.
CC HSP; P13727; I88U.
CC InterPro; IPR002352; Emaior_basic.
CC InterPro; IPR001304; Lectin_C.
CC Pfam; PF00059; lectin_c; 1.

```

```
DR PRINTS: PR00770; EMBL:AB000169; AAC73749.1; -
DR SMART: SM00034; CUECT; 1.
DR PROSITE: PS00615; C-TYPE LECTIN 1; 1.
DR PROSITE: PS00641; C-TYPE LECTIN 2; 1.
KW Eosinophil; Signal; Immune response; Antibiotic; Lectin;
KW Multigene family.
FT SIGNAL 1 15 POTENTIAL.
FT PROPEP 16 114 ACIDIC.
FT CHAIN 115 233 EOSINOPHIL GRANULE MAJOR BASIC PROTEIN 1.
FT DOMAIN 132 233 C-TYPE LECTIN (SHORT FORM).
FT DISULFID 134 231 BY SIMILARITY.
FT DISULFID 208 223 BY SIMILARITY.
SQ SEQUENCE 233 AA; 26268 MW; C8D5E96D927C56C8 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALL 1159
Db 4 LLLALL 10

RESULT 73
YBEO_ECOLI STANDARD; PRT; 235 AA.
AC P77427;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ybeu.
GN YBEO OR B0648.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,
RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map."
RL DNA Res. 3:1137-155(1996).
CC -!- SIMILARITY: STRONG, TO E.COLI YBER.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC or send an email to license@isb-sib.ch).
CC EMBL; U28377; AAA69134.1; -
DR EMBL; AE000379; AAC76003.1; -
DR PIR; E55082; E55082.
DR EcoGene; EG12987; yqga.
DR InterPro; IPR007563; DUF554.
DR Pfam; PF04474; DUF554; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 2 22 POTENTIAL.
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 56 76 POTENTIAL.
FT TRANSMEM 102 122 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 178 198 POTENTIAL.
FT TRANSMEM 210 230 POTENTIAL.
SQ SEQUENCE 235 AA; 24614 MW; CC89D2A93FED29EA CRC64;

Query Match 0.6%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 926 ELAAGSD 932
Db 202 ELAAGSD 208

RESULT 74
YQGA_ECOLI STANDARD; PRT; 235 AA.
AC Q46831;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yqga.
GN YQGA OR B2966.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Chung E., Allen E., Araujo R., Aparicio A., Davis K., Duncan M.,
RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map."
RL DNA Res. 3:1137-155(1996).
CC -!- SIMILARITY: STRONG, TO E.COLI YBER.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC or send an email to license@isb-sib.ch).
CC EMBL; U28377; AAA69134.1; -
DR EMBL; AE000379; AAC76003.1; -
DR PIR; E55082; E55082.
DR EcoGene; EG12987; yqga.
DR InterPro; IPR007563; DUF554.
DR Pfam; PF04474; DUF554; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 2 22 POTENTIAL.
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 56 76 POTENTIAL.
FT TRANSMEM 102 122 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 178 198 POTENTIAL.
FT TRANSMEM 210 230 POTENTIAL.
SQ SEQUENCE 235 AA; 24614 MW; CC89D2A93FED29EA CRC64;

Query Match 0.6%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1151 GGLLLA 1157
Db 194 GGLLLA 200

RESULT 75
RR2_MAIZE
ID_R2_MAIZE STANDARD; PRT; 236 AA.
AC P16037;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S2.
GN RPS2.
OS Zea mays (Maize).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90175001; PubMed=2308853;
RA Igloi G.L., Meinke A., Doery I., Koessel H.;
RT "Nucleotide and derived amino acid sequence of rps2 from maize
RT chloroplasts.";
RL Nucleic Acids Res. 18:663-663(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CV, FR9CMSR37; TISSUE=Leaf;
RX MEDLINE=90272437; PubMed=2140886;
RA Stahl D., Rodermeil S., Subramanian A.R., Bogorad L.;
RT "Nucleotide sequence of a 3.46 kb region of maize chloroplast DNA
RT containing the gene cluster rpoC2-rps2-atpI-atpH.";
RL Nucleic Acids Res. 18:3073-3074(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95395841; PubMed=7666415;
RA Maier R.M., Neckermann K., Igloi G.L., Koessel H.;
RT "Complete sequence of the maize chloroplast genome: gene content,
RT hotspots of divergence and fine tuning of genetic information by
RT transcript editing.";
RL J. Mol. Biol. 251:614-628(1995).
RN [4]
RP SEQUENCE OF 1-36 FROM N.A.
RX MEDLINE=90340289; PubMed=2381419;
RA Igloi G.L., Meinke A., Doery I., Koessel H.;
RT "Nucleotide sequence of the maize chloroplast rpo B/C1/C2 operon:
RT comparison between the derived protein primary structures from
RT various organisms with respect to functional domains.";
RL Mol. Gen. Genet. 221:379-394(1990).
CC -!- SIMILARITY: Belongs to the S2P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17318; CAA35198.1; -
DR EMBL; X52270; CAA36512.1; -
DR EMBL; X86563; CAA60279.1; -
DR FIR; S08249; B32M2.
DR Gramene; P16037; -
DR MaizeDB; 66014; -
DR HAMAP; MF_00291; -; 1.
DR InterPro; IPR001865; Ribosomal_S2.
DR InterPro; IPR005706; Ribosomal_S2_b/o.
DR Pfam; PF00318; Ribosomal_S2; 1.
DR PRINTS; PR00395; RIBOSOMALS2.
DR TIGRfam; TIGR01011; rpsB_bact; 1.
```

```
DR PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
DR PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
KW Ribosomal protein; Chloroplast.
SQ SEQUENCE 236 AA; 26881 MW; 80A3F99A8955FAD5 CRC64;

Query Match      0.6%; Score 7; DB 1; Length 236;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLD 170
Db 163 DIVIVLD 169

Search completed: June 24, 2004, 18:04:24
Job time : 23 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 18:01:31 ; Search time 57 Seconds  
(without alignments)  
6576.064 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 1188

Sequence: 1 MDLPRGLVVALSLWPGFT.....FRSARRRRRPGLDTPKVLK 1188

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

SPREMBL\_25.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phase.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriap.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	662	55.7	823	4 Q8WY18	Q8WY18 homo sapien
2	122	10.3	1188	11 Q7TQC3	Q7TQC3 mus musculus
3	45	3.8	823	11 Q8CE84	Q8CE84 mus musculus
4	15	1.3	1171	13 Q42094	Q42094 gallus gall
5	11	0.9	108	11 Q8BM12	Q8BM12 mus musculus
6	10	0.8	116	13 Q81670	Q81670 xenopus lae
7	10	0.8	399	16 Q8XSH5	Q8XSH5 raietonia s
8	9	0.8	164	4 Q8N112	Q8N112 homo sapien
9	9	0.8	315	16 Q8YCR3	Q8YCR3 brucella me
10	9	0.8	316	16 Q8FVL4	Q8FVL4 brucella su
11	9	0.8	400	6 Q8HY16	Q8HY16 cebus apell
12	9	0.8	413	16 Q89LS8	Q89LS8 bradyrhizob
13	9	0.8	457	13 Q73804	Q73804 fugu rubrip
14	9	0.8	607	10 Q39775	Q39775 gnetum gnet
15	9	0.8	780	13 Q62711	Q62711 xenopus lae
16	8	0.7	54	6 Q9TS65	Q9TS65 canis famil

Q88NW3 pseudomonas  
Q8CEU3 mus musculu  
Q9N241 saimiri bol  
Q9N242 ateles sp.  
Q9N243 homo sapien  
Q9N244 nasalis lar  
Q9N245 macaca mula  
Q9N246 papio hamad  
Q9N247 hylobates s  
Q9N248 hylobates k  
Q9N249 hylobates l  
Q9N250 gorilla gor  
Q9N251 pongo pygma  
Q9N252 pongo pygma  
Q9N253 pongo pygma  
Q9N254 pongo pygma  
Q9N255 pongo pygma  
Q9N256 pongo pygma  
Q9N257 pongo pygma  
Q9N258 pongo pygma  
Q9N259 pongo pygma  
Q9N260 pongo pygma  
Q9N261 pongo pygma  
Q9N262 pongo pygma  
Q9N263 pongo pygma  
Q9N264 pongo pygma  
Q9N265 pongo pygma  
Q9N266 pongo pygma  
Q9N267 pongo pygma  
Q9N268 pongo pygma  
Q9N269 pongo pygma  
Q9N270 pongo pygma  
Q9N271 pongo pygma  
Q9N272 pongo pygma  
Q9N273 pongo pygma  
Q9N274 pongo pygma  
Q9N275 pongo pygma  
Q9N276 pongo pygma  
Q9N277 pongo pygma  
Q9N278 pongo pygma  
Q9N279 pongo pygma  
Q9N280 pongo pygma  
Q9N281 pongo pygma  
Q9N282 pongo pygma  
Q9N283 pongo pygma  
Q9N284 pongo pygma  
Q9N285 pongo pygma  
Q9N286 pongo pygma  
Q9N287 pongo pygma  
Q9N288 pongo pygma  
Q9N289 pongo pygma  
Q9N290 pongo pygma  
Q9N291 pongo pygma  
Q9N292 pongo pygma  
Q9N293 pongo pygma  
Q9N294 pongo pygma  
Q9N295 pongo pygma  
Q9N296 pongo pygma  
Q9N297 pongo pygma  
Q9N298 pongo pygma  
Q9N299 pongo pygma  
Q9N300 pongo pygma  
Q9N301 pongo pygma  
Q9N302 pongo pygma  
Q9N303 pongo pygma  
Q9N304 pongo pygma  
Q9N305 pongo pygma  
Q9N306 pongo pygma  
Q9N307 pongo pygma  
Q9N308 pongo pygma  
Q9N309 pongo pygma  
Q9N310 pongo pygma  
Q9N311 pongo pygma  
Q9N312 pongo pygma  
Q9N313 pongo pygma  
Q9N314 pongo pygma  
Q9N315 pongo pygma  
Q9N316 pongo pygma  
Q9N317 pongo pygma  
Q9N318 pongo pygma  
Q9N319 pongo pygma  
Q9N320 pongo pygma  
Q9N321 pongo pygma  
Q9N322 pongo pygma  
Q9N323 pongo pygma  
Q9N324 pongo pygma  
Q9N325 pongo pygma  
Q9N326 pongo pygma  
Q9N327 pongo pygma  
Q9N328 pongo pygma  
Q9N329 pongo pygma  
Q9N330 pongo pygma  
Q9N331 pongo pygma  
Q9N332 pongo pygma  
Q9N333 pongo pygma  
Q9N334 pongo pygma  
Q9N335 pongo pygma  
Q9N336 pongo pygma  
Q9N337 pongo pygma  
Q9N338 pongo pygma  
Q9N339 pongo pygma  
Q9N340 pongo pygma  
Q9N341 pongo pygma  
Q9N342 pongo pygma  
Q9N343 pongo pygma  
Q9N344 pongo pygma  
Q9N345 pongo pygma  
Q9N346 pongo pygma  
Q9N347 pongo pygma  
Q9N348 pongo pygma  
Q9N349 pongo pygma  
Q9N350 pongo pygma  
Q9N351 pongo pygma  
Q9N352 pongo pygma  
Q9N353 pongo pygma  
Q9N354 pongo pygma  
Q9N355 pongo pygma  
Q9N356 pongo pygma  
Q9N357 pongo pygma  
Q9N358 pongo pygma  
Q9N359 pongo pygma  
Q9N360 pongo pygma  
Q9N361 pongo pygma  
Q9N362 pongo pygma  
Q9N363 pongo pygma  
Q9N364 pongo pygma  
Q9N365 pongo pygma  
Q9N366 pongo pygma  
Q9N367 pongo pygma  
Q9N368 pongo pygma  
Q9N369 pongo pygma  
Q9N370 pongo pygma  
Q9N371 pongo pygma  
Q9N372 pongo pygma  
Q9N373 pongo pygma  
Q9N374 pongo pygma  
Q9N375 pongo pygma  
Q9N376 pongo pygma  
Q9N377 pongo pygma  
Q9N378 pongo pygma  
Q9N379 pongo pygma  
Q9N380 pongo pygma  
Q9N381 pongo pygma  
Q9N382 pongo pygma  
Q9N383 pongo pygma  
Q9N384 pongo pygma  
Q9N385 pongo pygma  
Q9N386 pongo pygma  
Q9N387 pongo pygma  
Q9N388 pongo pygma  
Q9N389 pongo pygma  
Q9N390 pongo pygma  
Q9N391 pongo pygma  
Q9N392 pongo pygma  
Q9N393 pongo pygma  
Q9N394 pongo pygma  
Q9N395 pongo pygma  
Q9N396 pongo pygma  
Q9N397 pongo pygma  
Q9N398 pongo pygma  
Q9N399 pongo pygma  
Q9N400 pongo pygma  
Q9N401 pongo pygma  
Q9N402 pongo pygma  
Q9N403 pongo pygma  
Q9N404 pongo pygma  
Q9N405 pongo pygma  
Q9N406 pongo pygma  
Q9N407 pongo pygma  
Q9N408 pongo pygma  
Q9N409 pongo pygma  
Q9N410 pongo pygma  
Q9N411 pongo pygma  
Q9N412 pongo pygma  
Q9N413 pongo pygma  
Q9N414 pongo pygma  
Q9N415 pongo pygma  
Q9N416 pongo pygma  
Q9N417 pongo pygma  
Q9N418 pongo pygma  
Q9N419 pongo pygma  
Q9N420 pongo pygma  
Q9N421 pongo pygma  
Q9N422 pongo pygma  
Q9N423 pongo pygma  
Q9N424 pongo pygma  
Q9N425 pongo pygma  
Q9N426 pongo pygma  
Q9N427 pongo pygma  
Q9N428 pongo pygma  
Q9N429 pongo pygma  
Q9N430 pongo pygma  
Q9N431 pongo pygma  
Q9N432 pongo pygma  
Q9N433 pongo pygma  
Q9N434 pongo pygma  
Q9N435 pongo pygma  
Q9N436 pongo pygma  
Q9N437 pongo pygma  
Q9N438 pongo pygma  
Q9N439 pongo pygma  
Q9N440 pongo pygma  
Q9N441 pongo pygma  
Q9N442 pongo pygma  
Q9N443 pongo pygma  
Q9N444 pongo pygma  
Q9N445 pongo pygma  
Q9N446 pongo pygma  
Q9N447 pongo pygma  
Q9N448 pongo pygma  
Q9N449 pongo pygma  
Q9N450 pongo pygma  
Q9N451 pongo pygma  
Q9N452 pongo pygma  
Q9N453 pongo pygma  
Q9N454 pongo pygma  
Q9N455 pongo pygma  
Q9N456 pongo pygma  
Q9N457 pongo pygma  
Q9N458 pongo pygma  
Q9N459 pongo pygma  
Q9N460 pongo pygma  
Q9N461 pongo pygma  
Q9N462 pongo pygma  
Q9N463 pongo pygma  
Q9N464 pongo pygma  
Q9N465 pongo pygma  
Q9N466 pongo pygma  
Q9N467 pongo pygma  
Q9N468 pongo pygma  
Q9N469 pongo pygma  
Q9N470 pongo pygma  
Q9N471 pongo pygma  
Q9N472 pongo pygma  
Q9N473 pongo pygma  
Q9N474 pongo pygma  
Q9N475 pongo pygma  
Q9N476 pongo pygma  
Q9N477 pongo pygma  
Q9N478 pongo pygma  
Q9N479 pongo pygma  
Q9N480 pongo pygma  
Q9N481 pongo pygma  
Q9N482 pongo pygma  
Q9N483 pongo pygma  
Q9N484 pongo pygma  
Q9N485 pongo pygma  
Q9N486 pongo pygma  
Q9N487 pongo pygma  
Q9N488 pongo pygma  
Q9N489 pongo pygma  
Q9N490 pongo pygma  
Q9N491 pongo pygma  
Q9N492 pongo pygma  
Q9N493 pongo pygma  
Q9N494 pongo pygma  
Q9N495 pongo pygma  
Q9N496 pongo pygma  
Q9N497 pongo pygma  
Q9N498 pongo pygma  
Q9N499 pongo pygma  
Q9N500 pongo pygma  
Q9N501 pongo pygma  
Q9N502 pongo pygma  
Q9N503 pongo pygma  
Q9N504 pongo pygma  
Q9N505 pongo pygma  
Q9N506 pongo pygma  
Q9N507 pongo pygma  
Q9N508 pongo pygma  
Q9N509 pongo pygma  
Q9N510 pongo pygma  
Q9N511 pongo pygma  
Q9N512 pongo pygma  
Q9N513 pongo pygma  
Q9N514 pongo pygma  
Q9N515 pongo pygma  
Q9N516 pongo pygma  
Q9N517 pongo pygma  
Q9N518 pongo pygma  
Q9N519 pongo pygma  
Q9N520 pongo pygma  
Q9N521 pongo pygma  
Q9N522 pongo pygma  
Q9N523 pongo pygma  
Q9N524 pongo pygma  
Q9N525 pongo pygma  
Q9N526 pongo pygma  
Q9N527 pongo pygma  
Q9N528 pongo pygma  
Q9N529 pongo pygma  
Q9N530 pongo pygma  
Q9N531 pongo pygma  
Q9N532 pongo pygma  
Q9N533 pongo pygma  
Q9N534 pongo pygma  
Q9N535 pongo pygma  
Q9N536 pongo pygma  
Q9N537 pongo pygma  
Q9N538 pongo pygma  
Q9N539 pongo pygma  
Q9N540 pongo pygma  
Q9N541 pongo pygma  
Q9N542 pongo pygma  
Q9N543 pongo pygma  
Q9N544 pongo pygma  
Q9N545 pongo pygma  
Q9N546 pongo pygma  
Q9N547 pongo pygma  
Q9N548 pongo pygma  
Q9N549 pongo pygma  
Q9N550 pongo pygma  
Q9N551 pongo pygma  
Q9N552 pongo pygma  
Q9N553 pongo pygma  
Q9N554 pongo pygma  
Q9N555 pongo pygma  
Q9N556 pongo pygma  
Q9N557 pongo pygma  
Q9N558 pongo pygma  
Q9N559 pongo pygma  
Q9N560 pongo pygma  
Q9N561 pongo pygma  
Q9N562 pongo pygma  
Q9N563 pongo pygma  
Q9N564 pongo pygma  
Q9N565 pongo pygma  
Q9N566 pongo pygma  
Q9N567 pongo pygma  
Q9N568 pongo pygma  
Q9N569 pongo pygma  
Q9N570 pongo pygma  
Q9N571 pongo pygma  
Q9N572 pongo pygma  
Q9N573 pongo pygma  
Q9N574 pongo pygma  
Q9N575 pongo pygma  
Q9N576 pongo pygma  
Q9N577 pongo pygma  
Q9N578 pongo pygma  
Q9N579 pongo pygma  
Q9N580 pongo pygma  
Q9N581 pongo pygma  
Q9N582 pongo pygma  
Q9N583 pongo pygma  
Q9N584 pongo pygma  
Q9N585 pongo pygma  
Q9N586 pongo pygma  
Q9N587 pongo pygma  
Q9N588 pongo pygma  
Q9N589 pongo pygma  
Q9N590 pongo pygma  
Q9N591 pongo pygma  
Q9N592 pongo pygma  
Q9N593 pongo pygma  
Q9N594 pongo pygma  
Q9N595 pongo pygma  
Q9N596 pongo pygma  
Q9N597 pongo pygma  
Q9N598 pongo pygma  
Q9N599 pongo pygma  
Q9N600 pongo pygma  
Q9N601 pongo pygma  
Q9N602 pongo pygma  
Q9N603 pongo pygma  
Q9N604 pongo pygma  
Q9N605 pongo pygma  
Q9N606 pongo pygma  
Q9N607 pongo pygma  
Q9N608 pongo pygma  
Q9N609 pongo pygma  
Q9N610 pongo pygma  
Q9N611 pongo pygma  
Q9N612 pongo pygma  
Q9N613 pongo pygma  
Q9N614 pongo pygma  
Q9N615 pongo pygma  
Q9N616 pongo pygma  
Q9N617 pongo pygma  
Q9N618 pongo pygma  
Q9N619 pongo pygma  
Q9N620 pongo pygma  
Q9N621 pongo pygma  
Q9N622 pongo pygma  
Q9N623 pongo pygma  
Q9N624 pongo pygma  
Q9N625 pongo pygma  
Q9N626 pongo pygma  
Q9N627 pongo pygma  
Q9N628 pongo pygma  
Q9N629 pongo pygma  
Q9N630 pongo pygma  
Q9N631 pongo pygma  
Q9N632 pongo pygma  
Q9N633 pongo pygma  
Q9N634 pongo pygma  
Q9N635 pongo pygma  
Q9N636 pongo pygma  
Q9N637 pongo pygma  
Q9N638 pongo pygma  
Q9N639 pongo pygma  
Q9N640 pongo pygma  
Q9N641 pongo pygma  
Q9N642 pongo pygma  
Q9N643 pongo pygma  
Q9N644 pongo pygma  
Q9N645 pongo pygma  
Q9N646 pongo pygma  
Q9N647 pongo pygma  
Q9N648 pongo pygma  
Q9N649 pongo pygma  
Q9N650 pongo pygma  
Q9N651 pongo pygma  
Q9N652 pongo pygma  
Q9N653 pongo pygma  
Q9N654 pongo pygma  
Q9N655 pongo pygma  
Q9N656 pongo pygma  
Q9N657 pongo pygma  
Q9N658 pongo pygma  
Q9N659 pongo pygma  
Q9N660 pongo pygma  
Q9N661 pongo pygma  
Q9N662 pongo pygma  
Q9N663 pongo pygma  
Q9N664 pongo pygma  
Q9N665 pongo pygma  
Q9N666 pongo pygma  
Q9N667 pongo pygma  
Q9N668 pongo pygma  
Q9N669 pongo pygma  
Q9N670 pongo pygma  
Q9N671 pongo pygma  
Q9N672 pongo pygma  
Q9N673 pongo pygma  
Q9N674 pongo pygma  
Q9N675 pongo pygma  
Q9N676 pongo pygma  
Q9N677 pongo pygma  
Q9N678 pongo pygma  
Q9N679 pongo pygma  
Q9N680 pongo pygma  
Q9N681 pongo pygma  
Q9N682 pongo pygma  
Q9N683 pongo pygma  
Q9N684 pongo pygma  
Q9N685 pongo pygma  
Q9N686 pongo pygma  
Q9N687 pongo pygma  
Q9N688 pongo pygma  
Q9N689 pongo pygma  
Q9N690 pongo pygma  
Q9N691 pongo pygma  
Q9N692 pongo pygma  
Q9N693 pongo pygma  
Q9N694 pongo pygma  
Q9N695 pongo pygma  
Q9N696 pongo pygma  
Q9N697 pongo pygma  
Q9N698 pongo pygma  
Q9N699 pongo pygma  
Q9N700 pongo pygma  
Q9N701 pongo pygma  
Q9N702 pongo pygma  
Q9N703 pongo pygma  
Q9N704 pongo pygma  
Q9N705 pongo pygma  
Q9N706 pongo pygma  
Q9N707 pongo pygma  
Q9N708 pongo pygma  
Q9N709 pongo pygma  
Q9N710 pongo pygma  
Q9N711 pongo pygma  
Q9N712 pongo pygma  
Q9N713 pongo pygma  
Q9N714 pongo pygma  
Q9N715 pongo pygma  
Q9N716 pongo pygma  
Q9N717 pongo pygma  
Q9N718 pongo pygma  
Q9N719 pongo pygma  
Q9N720 pongo pygma  
Q9N721 pongo pygma  
Q9N722 pongo pygma  
Q9N723 pongo pygma  
Q9N724 pongo pygma  
Q9N725 pongo pygma  
Q9N726 pongo pygma  
Q9N727 pongo pygma  
Q9N728 pongo pygma  
Q9N729 pongo pygma  
Q9N730 pongo pygma  
Q9N731 pongo pygma  
Q9N732 pongo pygma  
Q9N733 pongo pygma  
Q9N734 pongo pygma  
Q9N735 pongo pygma  
Q9N736 pongo pygma  
Q9N737 pongo pygma  
Q9N738 pongo pygma  
Q9N739 pongo pygma  
Q9N740 pongo pygma  
Q9N741 pongo pygma  
Q9N742 pongo pygma  
Q9N743 pongo pygma  
Q9N744 pongo pygma  
Q9N745 pongo pygma  
Q9N746 pongo pygma  
Q9N747 pongo pygma  
Q9N748 pongo pygma  
Q9N749 pongo pygma  
Q9N750 pongo pygma  
Q9N751 pongo pygma  
Q9N752 pongo pygma  
Q9N753 pongo pygma  
Q9N754 pongo pygma  
Q9N755 pongo pygma  
Q9N756 pongo pygma  
Q9N757 pongo pygma  
Q9N758 pongo pygma  
Q9N759 pongo pygma  
Q9N760 pongo pygma  
Q9N761 pongo pygma  
Q9N762 pongo pygma  
Q9N763 pongo pygma  
Q9N764 pongo pygma  
Q9N765 pongo pygma  
Q9N766 pongo pygma  
Q9N767 pongo pygma  
Q9N768 pongo pygma  
Q9N769 pongo pygma  
Q9N770 pongo pygma  
Q9N771 pongo pygma  
Q9N772 pongo pygma  
Q9N773 pongo pygma  
Q9N774 pongo pygma  
Q9N775 pongo pygma  
Q9N776 pongo pygma  
Q9N777 pongo pygma  
Q9N778 pongo pygma  
Q9N779 pongo pygma  
Q9N780 pongo pygma  
Q9N781 pongo pygma  
Q9N782 pongo pygma  
Q9N783 pongo pygma  
Q9N784 pongo pygma  
Q9N785 pongo pygma  
Q9N786 pongo pygma  
Q9N787 pongo pygma  
Q9N788 pongo pygma  
Q9N789 pongo pygma  
Q9N790 pongo pygma  
Q9N791 pongo pygma  
Q9N792 pongo pygma  
Q9N793 pongo pygma  
Q9N794 pongo pygma  
Q9N795 pongo pygma  
Q9N796 pongo pygma  
Q9N797 pongo pygma  
Q9N798 pongo pygma  
Q9N799 pongo pygma  
Q9N800 pongo pygma  
Q9N801 pongo pygma  
Q9N802 pongo pygma  
Q9N803 pongo pygma  
Q9N804 pongo pygma  
Q9N805 pongo pygma  
Q9N806 pongo pygma  
Q9N807 pongo pygma  
Q9N808 pongo pygma  
Q9N809 pongo pygma  
Q9N810 pongo pygma  
Q9N811 pongo pygma  
Q9N812 pongo pygma  
Q9N813 pongo pygma  
Q9N814 pongo pygma  
Q9N815 pongo pygma  
Q9N816 pongo pygma  
Q9N817 pongo pygma  
Q9N818 pongo pygma  
Q9N819 pongo pygma  
Q9N820 pongo pygma  
Q9N821 pongo pygma  
Q9N822 pongo pygma  
Q9N823 pongo pygma  
Q9N824 pongo pygma  
Q9N825 pongo pygma  
Q9N826 pongo pygma  
Q9N827 pongo pygma  
Q9N828 pongo pygma  
Q9N829 pongo pygma  
Q9N830 pongo pygma  
Q9N831 pongo pygma  
Q9N832 pongo pygma  
Q9N833 pongo pygma  
Q9N834 pongo pygma  
Q9N835 pongo pygma  
Q9N836 pongo pygma  
Q9N837 pongo pygma  
Q9N838 pongo pygma  
Q9N839 pongo pygma  
Q9N840 pongo pygma  
Q9N841 pongo pygma  
Q9N842 pongo pygma  
Q9N843 pongo pygma  
Q9N844 pongo pygma  
Q9N845 pongo pygma  
Q9N846 pongo pygma  
Q9N847 pongo pygma  
Q9N848 pongo pygma  
Q9N849 pongo pygma  
Q9N850 pongo pygma  
Q9N851 pongo pygma  
Q9N852 pongo pygma  
Q9N853 pongo pygma  
Q9N854 pongo pygma  
Q9N855 pongo pygma  
Q9N856 pongo pygma  
Q9N857 pongo pygma  
Q9N858 pongo pygma  
Q9N859 pongo pygma  
Q9N860 pongo pygma  
Q9N861 pongo pygma  
Q9N862 pongo pygma  
Q9N863 pongo pygma  
Q9N864 pongo pygma  
Q9N865 pongo pygma  
Q9N866 pongo pygma  
Q9N867 pongo pygma  
Q9N868 pongo pygma  
Q9N869 pongo pygma  
Q9N870 pongo pygma  
Q9N871 pongo pygma  
Q9N872 pongo pygma  
Q9N873 pongo pygma  
Q9N874 pongo pygma  
Q9N875 pongo pygma  
Q9N876 pongo pygma  
Q9N877 pongo pygma  
Q9N878 pongo pygma  
Q9N879 pongo pygma  
Q9N880 pongo pygma  
Q9N881 pongo pygma  
Q9N882 pongo pygma  
Q9N883 pongo pygma  
Q9N884 pongo pygma  
Q9N885 pongo pygma  
Q9N886 pongo pygma  
Q9N887 pongo pygma  
Q9N888 pongo pygma  
Q9N889 pongo pygma  
Q9N890 pongo pygma  
Q9N891 pongo pygma  
Q9N892 pongo pygma  
Q9N893 pongo pygma  
Q9N894 pongo pygma  
Q9N895 pongo pygma  
Q9N896 pongo pygma  
Q9N897 pongo pygma  
Q9N898 pongo pygma  
Q9N899 pongo pygma  
Q9N900 pongo pygma  
Q9N901 pongo pygma  
Q9N902 pongo pygma  
Q9N903 pongo pygma  
Q9N904 pongo pygma  
Q9N905 pongo pygma  
Q9N906 pongo pygma  
Q9N907 pongo pygma  
Q9N908 pongo pygma  
Q9N909 pongo pygma  
Q9N910 pongo pygma  
Q9N911 pongo pygma  
Q9N912 pongo pygma  
Q9N913 pongo pygma  
Q9N914 pongo pygma  
Q9N915 pongo pygma  
Q9N916 pongo pygma  
Q9N917 pongo pygma  
Q9N918 pongo pygma  
Q9N919 pongo pygma  
Q9N920 pongo pygma  
Q9N921 pongo pygma  
Q9N922 pongo pygma  
Q9N923 pongo pygma  
Q9N924 pongo pygma  
Q9N925 pongo pygma  
Q9N926 pongo pygma  
Q9N927 pongo pygma  
Q9N928 pongo pygma  
Q9N929 pongo pygma  
Q9N930 pongo pygma  
Q9N931 pongo pygma  
Q9N932 pongo pygma  
Q9N933 pongo pygma  
Q9N934 pongo pygma  
Q9N935 pongo pygma  
Q9N936 pongo pygma  
Q9N937 pongo pygma  
Q9N938 pongo pygma  
Q9N939 pongo pygma  
Q9N940 pongo pygma  
Q9N941 pongo pygma  
Q9N942 pongo pygma  
Q9N943 pongo pygma  
Q9N944 pongo pygma  
Q9N945 pongo pygma  
Q9N946 pongo pygma  
Q9N947 pongo pygma  
Q9N948 pongo pygma  
Q9N949 pongo pygma  
Q9N950 pongo pygma  
Q9N951 pongo pygma  
Q9N952 pongo pygma  
Q9N953 pongo pygma  
Q9N954 pongo pygma  
Q9N955 pongo pygma  
Q9N956 pongo pygma  
Q9N957 pongo pygma  
Q9N958 pongo pygma  
Q9N959 pongo pygma  
Q9N960 pongo pygma  
Q9N961 pongo pygma  
Q9N962 pongo pygma  
Q9N963 pongo pygma  
Q9N964 pongo pygma  
Q9N965 pongo pygma  
Q9N966 pongo pygma  
Q9N967 p

90 Q8PG60 xanthomonas 8 0.7 422 16 Q8PG60  
91 Q9P3U1 schizosacch 8 0.7 434 3 Q9P3U1  
92 Q8YE57 brucella me 8 0.7 442 16 Q8YE57  
93 Q8FX8 brucella su 8 0.7 442 16 Q8FX8  
94 Q04611 arabidopsis 8 0.7 444 10 Q04611  
95 Q8EWH6 mycoplasma 8 0.7 444 16 Q8EWH6  
96 Q9EYG0 amycolatops 8 0.7 463 2 Q9EYG0  
97 Q821Z0 streptomyc 8 0.7 484 16 Q821Z0  
98 Q52642 pseudomonas 8 0.7 487 2 Q52642  
99 Q51991 pseudomonas 8 0.7 491 2 Q51991  
100 Q64584 rattus norv 8 0.7 491 11 Q64584

# ALIGNMENTS

## RESULT 1

Q8WY18 PRELIMINARY; PRT; 823 AA.  
AC Q8WY18  
DT 01-MAR-2002 (T-EMBLrel. 20, Created)  
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)  
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)  
DE XSTP018.  
GN MST018.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Aorta;  
RA Xu Y.Y., Sun L.Z., Wu Q.Y., Liu Y.Q., Liu B., Zhao B., Wang X.Y.,  
RA Song L., Ye J., Sheng H., Gao Y., Zhang C.L., Zhang J., Wei Y.J.,  
RA Sun Y.H., Jiang Y.X., Zhao X.W., Liu S., Liu L.S., Ding J.P.,  
RA Gao R.L., Qiang B.Q., Yuan J.G., Liew C.C., Zhao M.S., Hui R.T.;  
RA Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
RL EMBL: AF111799; AAC139001.1;  
DR GO: 0008305; C: integrin complex; IEA.  
DR GO: 0004895; F: cell adhesion receptor activity; IEA.  
DR GO: 0007160; P: cell-matrix adhesion; IEA.  
DR InterPro: IPR000413; Integrin\_alpha.  
DR Pfam: PF01839; FG-GAP; 3.  
DR PRINTS: PR01185; INTEGRIN.  
DR SMART: SM00191; Int. alpha; 4.  
SQ SEQUENCE 823 AA; 92672 MW; DE4E78079DCD4925 CRC64;

Query Match 55.7%; Score 662; DB 4; Length 823;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 762; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 366 MSQTGFSSHVDEGVLLGAVGAYDNGAVLKTSGAVIPLRESYLKEPPEELKNGAYL 425  
Db 1 MSQTGFSSHVDEGVLLGAVGAYDNGAVLKTSGAVIPLRESYLKEPPEELKNGAYL 60  
Qy 426 GYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMNNRSLTIHQAMRGQIGSYFGSEIT 485  
Db 61 GYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMNNRSLTIHQAMRGQIGSYFGSEIT 120  
Qy 486 SVIDDGVTDVLLVGPMPYNEGRGKVVYELRQNVYNGTLKDSHYONARFGSS 545  
Db 121 SVIDDGVTDVLLVGPMPYNEGRGKVVYELRQNVYNGTLKDSHYONARFGSS 180  
Qy 546 IASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGRGSIILKTPKQRTASELATGLQYFG 605  
Db 181 IASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGRGSIILKTPKQRTASELATGLQYFG 240  
Qy 606 CSIHGQDLNEDGLIDLAVGALGNVILNSRPVQVQINASLHFPSPKINIFHRCKSGRD 665  
Db 241 CSIHGQDLNEDGLIDLAVGALGNVILNSRPVQVQINASLHFPSPKINIFHRCKSGRD 300  
Qy 666 ATCLAAFLCPTPIFLAPHFOTTTVGIRYNATMDERRYTPRAHLDEGGDRFTNRAVLSSG 725

301 ATCLAAFLCPTPIFLAPHFOTTTVGIRYNATMDERRYTPRAHLDEGGDRFTNRAVLSSG 360  
Qy 726 QELCERINPHVLDADYVVKPVTFESVEYSLEDPDHGPMDDGMPPTLRVSVFPWNGCNDE 785  
Db 361 QELCERINPHVLDADYVVKPVTFESVEYSLEDPDHGPMDDGMPPTLRVSVFPWNGCNDE 420  
Qy 786 HCVPLDLVLDARSDLPAMBYCORVLRKPAQCSAYTLSFDTTVFIIESTRQVAVEATLE 845  
Db 421 HCVPLDLVLDARSDLPAMBYCORVLRKPAQCSAYTLSFDTTVFIIESTRQVAVEATLE 480  
Qy 846 NRGENAYTVLNISOSANLOFASLIQKEDSGSTECVNEERLQKQVNSVPPFRAKAK 905  
Db 481 NRGENAYTVLNISOSANLOFASLIQKEDSGSTECVNEERLQKQVNSVPPFRAKAK 540  
Qy 906 VAFRLDSEFSKSIPLHLEIEIAGSDNSNERDSTKEDNVAPLRPHLKYEADVLFTRSSSL 965  
Db 541 VAFRLDSEFSKSIPLHLEIEIAGSDNSNERDSTKEDNVAPLRPHLKYEADVLFTRSSSL 600  
Qy 966 SHYEVKLASSLERYDGIQPPFSCIFRIQNLGLPIHGMMKTIPIATRSNELLKLRDP 1025  
Db 601 SHYEVKLASSLERYDGIQPPFSCIFRIQNLGLPIHGMMKTIPIATRSNELLKLRDP 660  
Qy 1026 LTDEANTSCNIWGNSTERYPTVEEDLRAPQLAHNSDVVSNICNIRLVPNQEINPHLL 1085  
Db 661 LTDEANTSCNIWGNSTERYPTVEEDLRAPQLAHNSDVVSNICNIRLVPNQEINPHLL 720  
Qy 1086 GNLWLRSLKALKYKSMKIMVNAALQRFHSPFIREDPSRQI 1128  
Db 721 GNLWLRSLKALKYKSMKIMVNAALQRFHSPFIREDPSRQI 763

## RESULT 2

Q7TQC3 PRELIMINARY; PRT; 1188 AA.  
AC Q7TQC3  
DT 01-OCT-2003 (T-EMBLrel. 25, Created)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
DE All integrin.  
GN ITGALL.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Tiger C.-F., Popova S.N., Velling T., Ekman S., Forsberg E.,  
RA Gullberg D.;  
RT "allb1 integrin is important for mesenchymal cell function:  
RT elimination of allb1 leads to dwarfism."  
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Johansson M., Popova S.N.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY124460; AAM62130.1; -.  
KW Integrin.  
SQ SEQUENCE 1188 AA; 133067 MW; 69B2BF20D75E4271 CRC64;

Query Match 10.3%; Score 122; DB 11; Length 1188;  
Best Local Similarity 100.0%; Pred. No. 8.3e-118;  
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 81 GNCTKLNLRVTLNSVSRKONMRLGLSLATNPKNDSFLACPLWSHBCGSSYYTGMCS 140  
Db 81 GNCTKLNLRVTLNSVSRKONMRLGLSLATNPKNDSFLACPLWSHBCGSSYYTGMCS 140  
Qy 141 RVNSFRSKTVAPALQRCOTMDIVIVLDGNSIYPWVEQVHFLINILKXFIYIGPQIQ 200  
Db 141 RVNSFRSKTVAPALQRCOTMDIVIVLDGNSIYPWVEQVHFLINILKXFIYIGPQIQ 200  
Qy 201 VG 202

DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR Pfam; PF00092; vwa; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR00453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 4.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS00234; VWF\_A; 1.  
KW Integrin.  
SQ SEQUENCE 1171 AA; 130228 MW; B505A4D65F09736E CRC64;  
  
Query Match 1.3%; Score 15; DB 13; Length 1171;  
Best Local Similarity 100.0%; Pred.No. 8.8e-06; Mismatches 0; Indels 0; Gaps 0;  
Matches 15; Conservative 0;  
  
Qy 164 DIVIVDGSNSIYFW 178  
Db 162 DIVIVDGSNSIYFW 176  
|||||  
|||  
  
RESULT 5  
Q8BM12 PRELIMINARY; PRT; 288 AA.  
ID Q8BM12  
AC Q8BM12  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Integrin alpha-10 precursor homolog (Fragment).  
GN ITGAL0.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CS7BL/6J; TISSUE=Skin;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium.  
RT \*Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573 (2002).  
DR EMBL; AK037222; BAC29761.1; -.  
DR MGD; MGI:2153482; Itgal0.  
DR GO; GO:0008305; C:integrin complex; IEA.  
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.  
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
FT NON TER 1  
SQ SEQUENCE 288 AA; 31949 MW; 3F6200F9C9475BE5 CRC64;  
  
Query Match 0.9%; Score 11; DB 11; Length 288;  
Best Local Similarity 100.0%; Pred.No. 0.04; Mismatches 0; Indels 0; Gaps 0;  
Matches 11; Conservative 0;  
  
Qy 1150 LGGLLLALLV 1160  
Db 252 LGGLLLALLV 262  
|||||  
|||  
  
RESULT 6  
Q91670 PRELIMINARY; PRT; 116 AA.  
ID Q91670  
AC Q91670  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.  
DR GO; GO:0008305; C:integrin complex; IEA.  
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
FT NON TER 1  
SQ SEQUENCE 288 AA; 31949 MW; 3F6200F9C9475BE5 CRC64;  
  
Query Match 0.9%; Score 11; DB 11; Length 288;  
Best Local Similarity 100.0%; Pred.No. 0.04; Mismatches 0; Indels 0; Gaps 0;  
Matches 11; Conservative 0;  
  
Qy 1150 LGGLLLALLV 1160  
Db 252 LGGLLLALLV 262  
|||||  
|||  
  
RESULT 6  
Q91670 PRELIMINARY; PRT; 116 AA.  
ID Q91670  
AC Q91670  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.  
DR GO; GO:0008305; C:integrin complex; IEA.  
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR Pfam; PF00357; Integrin\_A; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
FT NON TER 1  
SQ SEQUENCE 288 AA; 31949 MW; 3F6200F9C9475BE5 CRC64;  
  
Query Match 0.9%; Score 11; DB 11; Length 288;  
Best Local Similarity 100.0%; Pred.No. 0.04; Mismatches 0; Indels 0; Gaps 0;  
Matches 11; Conservative 0;  
  
Qy 1150 LGGLLLALLV 1160  
Db 252 LGGLLLALLV 262  
|||||  
|||  
  
RESULT 6  
Q91670 PRELIMINARY; PRT; 116 AA.  
ID Q91670  
AC Q91670  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

```

01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Integrin alpha-1 (Fragment).
GN GENE 4.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312514; PubMed=8700860;
RA Brown D.D., Wang Z., Furlow J.D., Kanamori A., Schwartzman R.A.,
RA Remo B.F., Binder A.;
RT "The thyroid hormone-induced tail resorption program during Xenopus
RT laevis metamorphosis.";
RL EMBL; Natl. Acad. Sci. U.S.A. 93:1924-1929(1996).
DR EMBL; U44025; AAC59878.1; -.
DR GO; GO:0008305; C:Integrin complex; IEA.
DR GO; GO:0004895; F:Cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:Cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR SMART; SM00191; Int_alpha; 1.
FT NON TER 116 116
SQ SEQUENCE 116 AA; 12749 MW; 2A8A5CA3E83DD07E CRC64;

Query Match 0.8%; Score 10; DB 13; Length 116;
Best Local Similarity 100.0%; Pred. No. 0.2; Mismatches 0; Indels 0; Gaps 0;
Matches 10; Conservative 0;

Qy 39 AFFGYTVQQH 48
Db 39 AFFGYTVQQH 48
|||||
|

RESULT 7
Q8XSH5 ID Q8XSH5 PRELIMINARY; PRT; 399 AA.
AC Q8XSH5;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative transport transmembrane protein.
GN RSP0499 OR R500378.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OG Plasmid megaplasmid.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Axlatt M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646079; CAD17650.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0006457; P:protein folding; IEA.
DR InterPro; IPR001175; FKBP_PPase.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS00453; FKBP_PPase_1; 1.
DR PROSITE; PS50850; MFS; 1.
DR Plasmid; Complete proteome.
SQ SEQUENCE 399 AA; 42148 MW; 096BB97883DCA7D1 CRC64;

Query Match 0.8%; Score 10; DB 16; Length 399;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1152 GLLALLLV 1161
Db 161 GLLALLLV 170
|||||
|

RESULT 8
Q8N112 ID Q8N112 PRELIMINARY; PRT; 164 AA.
AC Q8N112;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-WAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical gene LOC132228 (Hypothetical protein FLJ38608).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA Nishi T., Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kikuchi H., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami K.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEO human cDNA sequencing project";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028000; AAH28000.1; -.
DR EMBL; AK095927; BAC04652.1; -.
KW Hypothetical protein.
SQ SEQUENCE 164 AA; 17868 MW; 803993458714252E CRC64;

Query Match 0.8%; Score 9; DB 4; Length 164;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 LLLALLLV 1161
Db 99 LLLALLLV 107
|||||
|

RESULT 9
Q8YCR3 ID Q8YCR3 PRELIMINARY; PRT; 315 AA.
AC Q8YCR3;
DT 01-WAR-2002 (TReMBLrel. 20, Created)
DT 01-WAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE PERMEASE.
GN EME110465.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G., Kaprat V., Redkar R.J., Patra G., Mujer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Ietesson J.-J.,
RA Haseikorn R., Kyrpides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";

```



Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).

RL DR EMBL; AB009683; AAL53707.1; -;  
DR PIR; AH3567; AH3567.  
DR InterPro; IPR002781; DUF81.  
DR Pfam; PF01925; DUF81; 1.  
KW Complete proteome.  
SQ SEQUENCE 315 AA; 33486 MW; 7A34836134FB2F7D CRC64;

Query Match 0.8%; Score 9; DB 16; Length 315;  
Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVLA 1162  
DB 283 LLLALLVLA 291

RESULT 10  
Q8FVL4 PRELIMINARY; PRT; 316 AA.  
AC Q8FVL4  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-JUN-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Membrane protein, putative.  
GN BRA0823.  
OS Brucella suis.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Brucellaceae; Brucella.  
OX NCBI\_TaxID=29461;  
RN [1]  
SQ SEQUENCE FROM N.A.  
STRAIN=1330 / Biovar 1;  
MEDLINE=2247741; PubMed=12271122;  
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,  
RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,  
RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,  
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Var Aken S.E.,  
RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,  
RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.,  
RT "The Brucella suis genome reveals fundamental similarities between  
animal and plant pathogens and symbionts."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).  
DR EMBL; AE014576; AAN33998.1; -;  
DR TIGR; BRA0823; -;  
DR InterPro; IPR002781; DUF81.  
DR Pfam; PF01925; DUF81; 1.  
KW Complete proteome.  
SQ SEQUENCE 316 AA; 33671 MW; 4DC0DC32D8AB8087 CRC64;

Query Match 0.8%; Score 9; DB 16; Length 316;  
Best Local Similarity 100.0%; Pred. No. 5.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVLA 1162  
DB 284 LLLALLVLA 292

RESULT 11  
Q8HYL6 PRELIMINARY; PRT; 400 AA.  
ID Q8HYL6  
AC Q8HYL6  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Poliovirus receptor.  
GN PVR.  
OS Cebus apella (Brown-capped capuchin).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Cebus.  
OX NCBI\_TaxID=9515;  
RN [1]  
SQ SEQUENCE 400 AA; 43322 MW; C6C3A59A0A285F3A CRC64;

Query Match 0.8%; Score 9; DB 16; Length 400;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVLA 1161  
DB 9 LLLALLVLA 17

RESULT 12  
Q89LS8 PRELIMINARY; PRT; 413 AA.  
ID Q89LS8  
AC Q89LS8  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Blr4465 protein.  
GN BLR4465.  
OS Bradyrhizobium japonicum.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Bradyrhizobium.  
OX NCBI\_TaxID=375;  
RN [1]  
SQ SEQUENCE FROM N.A.  
STRAIN=USDA 110;  
RX MEDLINE=22484998; PubMed=12597275;  
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,  
RA Sasamoto S., Watanabe A., Idegawa K., Iziguchi M., Kawashima K.,  
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,  
RA Tabata S.;  
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
Bradyrhizobium japonicum USDA110."  
RL DNA Res. 9:189-197(2002).  
DR EMBL; AF005951; BAC49730.1; -;  
DR GO; GO:0005489; P:electron transporter activity; IEA.  
DR GO; GO:0016811; P:hydrolase activity, acting on carbon-nitrog. .; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR004304; FmdA\_AmdA.  
DR Pfam; PF03069; FmdA\_AmdA; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
KW Complete proteome.  
SQ SEQUENCE 413 AA; 45159 MW; E2F309DCC2E8D82B CRC64;

Query Match 0.8%; Score 9; DB 16; Length 413;  
Best Local Similarity 100.0%; Pred. No. 6.8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 FSXTVAPAL 156  
Db 18 FSXTVAPAL 26

RESULT 13

O73804 O73804 PRELIMINARY; PRT; 497 AA.  
AC O73804;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorph; Tetraodontiformes;  
OC Tetraodontidae; Tetraodontidae; Takifugu.  
OX NCBI\_TaxID=31033;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=99177347; PubMed=10077531;  
RX Gellner K., Brenner S.;  
RT "Analysis of 143 kb of genomic DNA around the wnt1 locus of Fugu  
rubripes.";  
RL Genome Res. 9:251-258(1999).  
DR EMBL; AF056116; AAC34385.1; -.  
DR InterPro; IPR008075; Lipocalin1\_recep.  
DR InterPro; IPR006876; LMBR1.  
DR Pfam; PF04791; LMBR1; 1.  
DR PRINTS; PR01692; LIPOCALINR.  
KW Hypothetical protein.  
SQ SEQUENCE 497 AA; 56376 MW; 802743E98383B64 CRC64;

Query Match 0.8%; Score 9; DB 13; Length 497;  
Best Local Similarity 100.0%; Pred. No. 8;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVL 1161  
Db 169 LLLALLVL 177

RESULT 14

O39775 O39775 PRELIMINARY; PRT; 607 AA.  
AC O39775;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Legumin, 11S globulin.  
OS Gnetum gnemon (Bago).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Gnetales; Gnetales; Gnetales; Gnetales; Gnetales.  
OX NCBI\_TaxID=3382;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Seed;  
RX MEDLINE=9840765; PubMed=9767693;  
RA Shatov A.D., Braun H., Chesnokov Y.V., Horstmann C., Kakhovskaya I.A.,  
RA Baumlein H.;  
RT "Sequence peculiarity of Gnetalean legumin-like seed storage  
proteins.";  
RL J. Mol. Evol. 47:486-492(1998).  
DR EMBL; Z50779; CAA90642.1; -.  
DR PIR; S60658; S60658.  
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.  
DR InterPro; IPR006045; Cupin.  
DR Pfam; PF00190; Cupin; 2.  
DR PRINTS; PR00439; 11SGLOBULIN.  
SQ SEQUENCE 607 AA; 68910 MW; 7AD6DDC504FB3A60 CRC64;

Query Match 0.8%; Score 9; DB 10; Length 607;  
Best Local Similarity 100.0%; Pred. No. 9.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVL 1161  
Db 10 LLLALLVL 18

RESULT 15

Q06271 Q06271 PRELIMINARY; PRT; 780 AA.  
AC Q06271;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Integrin alpha-2 subunit (Fragment).  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
OC Xenopodidae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP Meng F., Desimone D.W.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 95-168 FROM N.A.  
RX MEDLINE=94008528; PubMed=8404528;  
RA Whittaker C.A., Desimone D.W.;  
RT "Integrin alpha subunit mRNAs are differentially expressed in early  
Xenopus embryos.";  
RL Development 117:1239-1249(1993).  
DR EMBL; L43058; AAA69770.1; -.  
DR PIR; I51524; I51524.  
DR GO; GO:0008305; C:integrin complex; IEA.  
DR GO; GO:0004895; P:cell adhesion receptor activity; IEA.  
DR GO; GO:0007150; P:cell-matrix adhesion; IEA.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR Pfam; PF01839; FG-GAP; 3.  
DR PRINTS; PR01185; INTEGRIN\_A; 1.  
DR SMART; SM00191; Int\_alpha; 4.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
FT NON\_TER 1  
SQ SEQUENCE 780 AA; 87017 MW; 9518B18C2B6BF637 CRC64;

Query Match 0.8%; Score 9; DB 13; Length 780;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGP 503  
Db 108 TDVLLVGP 116

RESULT 16

O9TS65 O9TS65 PRELIMINARY; PRT; 54 AA.  
AC O9TS65;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE LEUKOINTEGRIN alpha D beta 2 (Fragments).  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RN SEQUENCE.  
RP MEDLINE=95325609; PubMed=7541420;  
RX Danilenko D.M., Rossitto P.V., Van der Vieren M., Le Trong H.,

RA McDonough S.P., Affolter V.X., Moore P.F.;  
RT "A novel canine leukointegrin, alpha d beta 2, is expressed by  
RT specific macrophage subpopulations in tissue and a minor CD8+  
RT lymphocyte subpopulation in peripheral blood.";  
RL J. Immunol. 155:35-44(1995).  
FT NON TER 1  
PT NON CONS 23 24  
PT NON TER 54 54  
SQ SEQUENCE 54 AA; 5624 MW; C2FC9C2C2FDEC3B6C CRC64;

Query Match 0.7%; Score 8; DB 6; Length 54;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 LVVGAPLE 63  
Db 24 LVVGAPLE 31

RESULT 17  
Q88NW3 PRELIMINARY; PRT; 58 AA.  
AC Q88NW3;  
DT 01-JUN-2003 (TRENBLrel. 24, Created)  
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Conserved hypothetical protein.  
GN PP1091.  
OS Pseudomonas putida (strain KT2440).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=160488;  
RN [1]

SEQUENCE FROM N.A.  
RP MEDLINE=22423060; PubMed=12534463;  
RX Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
RA Martins dos Santos V.A.P., Fouts D.E., Gall S.R., Pop M., Holmes M.,  
RA Brinkac L., Nelson M., DeBoy R.T., Daugherty S., Kolonay J.,  
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,  
RA Chatterjee P., Holtzapple E., Scanlan D., Tran K., Moazzes A.,  
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,  
RA Lauber J., Stjepandic D., Hohlseisel J., Straetz M., Heim S.,  
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,  
RA Fraser C.M.;  
RT "Complete genome sequence and comparative analysis of the  
RT metabolically versatile Pseudomonas putida KT2440.";  
RL Environ. Microbiol. 4:799-808(2002).  
DR EMBL; A3C16778; AAN66716.1; --  
DR TIGR; PP1091; --  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 58 AA; 6734 MW; CD9350CE793D269 CRC64;

Query Match 0.7%; Score 3; DB 16; Length 58;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 884 EERLQKQ 891  
Db 43 EERLQKQ 50

RESULT 18  
Q8CEU3 PRELIMINARY; PRT; 76 AA.  
AC Q8CEU3;  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE CD24a antigen.  
GN CD24A.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Query Match 0.7%; Score 8; DB 6; Length 54;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 LVVGAPLE 63  
Db 24 LVVGAPLE 31

OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CS7BL/6J; TISSUE=Head;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RT Nature 420:563-573(2002).  
DR EMBL; AK014279; BAC25429.1; --  
DR FIK; FT0721; FT0721.  
DR MGD; MGI:98323; Cd24a.  
SQ SEQUENCE 76 AA; 7787 MW; 6853F12240F91AEB CRC64;

Query Match 0.7%; Score 8; DB 11; Length 76;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLALL 1159  
Db 12 GLLALL 19

RESULT 19  
Q9N241 PRELIMINARY; PRT; 78 AA.  
ID Q9N241;  
AC Q9N241;  
DT 01-OCT-2000 (TRENBLrel. 15, Created)  
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Saimiri boliviensis (Bolivian squirrel monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.  
OX NCBI\_TaxID=27679;  
RN [2]

SEQUENCE FROM N.A.  
RP MEDLINE=20215006; PubMed=10750051;  
RX Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
RT primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125670; AAF64385.1; --  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm.1; 1.  
DR PRINTS; PR00237; GPCR\_Rhodopsn.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON TER 78 78  
SQ SEQUENCE 78 AA; 7607 MW; D4E684AD6DF6375B CRC64;

Query Match 0.7%; Score 8; DB 6; Length 78;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALL 1161  
Db 64 LLLALL 71

RESULT 20  
Q9N242 PRELIMINARY; PRT; 87 AA.  
ID Q9N242;  
AC Q9N242;  
DT 01-OCT-2000 (TRENBLrel. 15, Created)  
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)

Query Match 0.7%; Score 8; DB 6; Length 78;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALL 1161  
Db 64 LLLALL 71

DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Ateles sp. (Spider monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateleinae; Ateles.  
OX NCBI\_TaxID=9511;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
RT Primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125669; AAF64384.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCRHHODOPSN.  
DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON TER 87 87  
SQ SEQUENCE 87 AA; 8388 MW; 64574359DAC555D4 CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 87;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
Db 73 LLLALLVL 80  
|||||  
  
RESULT 21  
ID Q99587 PRELIMINARY; PRT; 88 AA.  
AC Q99587;  
DT 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96362071; PubMed=8746407;  
RA Cichon S., Nothen M.M., Catalano M., Di Bella D., Maier W.,  
RA Lichtermann D., Minges J., Albus M., Borrmann M., Franzeck E., et al;  
RT "Identification of two novel polymorphisms and a rare deletion variant  
RT in the human dopamine D4 receptor gene.";  
RL Psychiatr. Genet. 5:97-103(1995).  
DR EMBL; S82918; AAB46803.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCRHHODOPSN.  
DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON TER 88 88  
SQ SEQUENCE 88 AA; 8482 MW; 314093AB4678CEDF CRC64;  
  
Query Match 0.7%; Score 8; DB 4; Length 88;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161

Db 74 LLLALLVL 81  
|||||  
  
RESULT 22  
Q9N244 PRELIMINARY; PRT; 91 AA.  
ID Q9N244;  
AC Q9N244;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
DR DRD4.  
GN DRD4.  
OS Nasalis larvatus (Proboscis monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Colobinae;  
OC Nasalis.  
OX NCBI\_TaxID=43780;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
RT Primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125667; AAF64382.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCRHHODOPSN.  
DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON TER 91 91  
SQ SEQUENCE 91 AA; 8827 MW; 67B3D0764BD439C CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 91;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
Db 77 LLLALLVL 84  
|||||  
  
RESULT 23  
Q9N245 PRELIMINARY; PRT; 91 AA.  
ID Q9N245;  
AC Q9N245;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
DR DRD4.  
GN DRD4.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  
OC Cercopitheinae; Macaca.  
OX NCBI\_TaxID=9544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
RT Primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125666; AAF64381.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.

```

DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 91
SQ SEQUENCE 91 AA; 8728 MW; 7CF86CAC7E0CF39C CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 6; Length 91;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 77 LLLALLVL 84

RESULT 24
O97517 PRELIMINARY; PRT; 91 AA.
AC O97517;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates."
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010297; AAC67225.1; -.
GO; GO:0016021; C:integral to membrane; IEA.
GO; GO:0004872; F:receptor activity; IEA.
GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 91
SQ SEQUENCE 91 AA; 8763 MW; 765DBF6B3E386E299 CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 6; Length 91;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 77 LLLALLVL 84

RESULT 26
O97521 PRELIMINARY; PRT; 91 AA.
AC O97521;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates."
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010301; AAC67229.1; -.
GO; GO:0016021; C:integral to membrane; IEA.
GO; GO:0004872; F:receptor activity; IEA.
GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 91
SQ SEQUENCE 91 AA; 8934 MW; 93CD0F1CFC8E8E2C CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 6; Length 91;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 77 LLLALLVL 84

RESULT 27
O9N243 PRELIMINARY; PRT; 91 AA.
ID O9N243

```

AC Q9N243; 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 25, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
DRD4.  
GN Presbytis cristata (Silvered langur).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;  
OC Presbytis.  
OX NCBI\_TaxID=36232;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
RT Primates";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125668; AAF64383.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PRINTS; PR00237; GPCRHHODPSN.  
DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER  
SQ SEQUENCE 91 AA; 8927 MW; 67E53D0764BD439C CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 91;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 77 LLLALLVL 84  
  
RESULT 28  
ID Q9N246 PRELIMINARY; PRT; 91 AA.  
AC Q9N246;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
DRD4.  
GN Papio hamadryas (Hamadryas baboon).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Papio.  
OX NCBI\_TaxID=9557;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
RT Primates";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125665; AAF64380.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PRINTS; PR00237; GPCRHHODPSN.  
DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER  
SQ SEQUENCE 91 AA; 8770 MW; 608F90C00E10839C CRC64;

QY 1154 LLLALLVL 1161  
DB 77 LLLALLVL 84  
  
RESULT 29  
ID Q9N247 PRELIMINARY; PRT; 93 AA.  
AC Q9N247;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
DRD4.  
GN Hylobates syndactylus (Siamang) (Symphalangus syndactylus).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.  
OX NCBI\_TaxID=9590;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
RT Primates";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125664; AAF64379.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PRINTS; PR00237; GPCRHHODPSN.  
DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER  
SQ SEQUENCE 93 AA; 8904 MW; 5D07B793F2B624E4 CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 93;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 79 LLLALLVL 86  
  
RESULT 30  
ID Q99586 PRELIMINARY; PRT; 95 AA.  
AC Q99586;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96362071; PubMed=8746407;  
RA Cichon S., Nothen M.M., Catalano M., Di Bella D., Maier W.,  
RA Lichtermann D., Minges J., Albus M., Borrmann M., Franzek B., et al;  
RT "Identification of two novel polymorphisms and a rare deletion variant  
RT in the human dopamine D4 receptor gene";  
RL Psychiatr. Genet. 5:97-103(1995).  
DR EMBL; S82917; AAB46802.1; -.

DR GO: 0004952; F:dopamine receptor activity; NAS.  
DR GO: 0007212; P:dopamine receptor signaling pathway; NAS.  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm 1; 1.  
DR PRINTS: PR00237; GPCR\_Rhodopsn.  
DR PROSITE: PS00262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER 95 95  
SQ SEQUENCE 95 AA; 9148 MW; 7163AF13B6156686 CRC64;  
  
Query Match 0.7%; Score 8; DB 4; Length 95;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88  
  
RESULT 31  
O97514 PRELIMINARY; PRT; 95 AA.  
AC O97514;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Pan troglodytes (Chimpanzee).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
OX NCBI\_TaxID=9598;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF010294; AAC67222.1; -.  
GO: 00016021; C:integral to membrane; IEA.  
GO: 0004872; F:receptor activity; IEA.  
GO: 0001584; P:rhodopsin-like receptor activity; IEA.  
GO: 0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm 1; 1.  
DR PRINTS: PR00237; GPCR\_Rhodopsn.  
DR PROSITE: PS00262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
FT NON\_TER 95 95  
SQ SEQUENCE 95 AA; 9119 MW; 02BD95057095863C CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 95;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88  
  
RESULT 32  
Q9N249 PRELIMINARY; PRT; 95 AA.  
AC Q9N249;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Hylobates klossii (Kloss's gibbon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.  
OX NCBI\_TaxID=9598;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF010294; AAC67222.1; -.  
GO: 00016021; C:integral to membrane; IEA.  
GO: 0004872; F:receptor activity; IEA.  
GO: 0001584; P:rhodopsin-like receptor activity; IEA.  
GO: 0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm 1; 1.  
DR PRINTS: PR00237; GPCR\_Rhodopsn.  
DR PROSITE: PS00262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
FT NON\_TER 95 95  
SQ SEQUENCE 95 AA; 9119 MW; 02BD95057095863C CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 95;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88  
  
RESULT 33  
Q9N248 PRELIMINARY; PRT; 95 AA.  
AC Q9N248;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Hylobates lar (Common gibbon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.  
OX NCBI\_TaxID=9580;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125663; AAF64378.1; -.  
GO: 00016021; C:integral to membrane; IEA.  
GO: 0004872; F:receptor activity; IEA.  
GO: 0001584; P:rhodopsin-like receptor activity; IEA.  
GO: 0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm 1; 1.  
DR PRINTS: PR00237; GPCR\_Rhodopsn.  
DR PROSITE: PS00262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
FT NON\_TER 95 95  
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 95;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88  
  
RESULT 34

OX NCBI\_TaxID=9587;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125662; AAF64377.1; -.  
GO: 00016021; C:integral to membrane; IEA.  
GO: 0004872; F:receptor activity; IEA.  
GO: 0001584; P:rhodopsin-like receptor activity; IEA.  
GO: 0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm 1; 1.  
DR PRINTS: PR00237; GPCR\_Rhodopsn.  
DR PROSITE: PS00262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER 95 95  
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 95;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88  
  
RESULT 33  
Q9N248 PRELIMINARY; PRT; 95 AA.  
AC Q9N248;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Dopamine D4 receptor (Fragment).  
GN DRD4.  
OS Hylobates lar (Common gibbon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.  
OX NCBI\_TaxID=9580;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20215006; PubMed=10750051;  
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;  
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in  
primates.";  
RL J. Exp. Zool. 288:32-38(2000).  
DR EMBL; AF125663; AAF64378.1; -.  
GO: 00016021; C:integral to membrane; IEA.  
GO: 0004872; F:receptor activity; IEA.  
GO: 0001584; P:rhodopsin-like receptor activity; IEA.  
GO: 0007186; P:G-protein coupled receptor protein signalin. .; IEA.  
DR InterPro: IPR000276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm 1; 1.  
DR PRINTS: PR00237; GPCR\_Rhodopsn.  
DR PROSITE: PS00262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON\_TER 95 95  
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;  
  
Query Match 0.7%; Score 8; DB 6; Length 95;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1154 LLLALLVL 1161  
DB 81 LLLALLVL 88  
  
RESULT 34

```

O97516
ID O97516 PRELIMINARY; PRT; 95 AA.
AC O97516;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deindard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010296; AAC67224.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . ; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
FT NON TER
SQ SEQUENCE 95 AA; 9121 MW; ECF6C50B3A809D33 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88
|||||

O97520
ID O97520 PRELIMINARY; PRT; 95 AA.
AC O97520;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deindard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010300; AAC67228.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . ; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
FT NON TER
SQ SEQUENCE 95 AA; 9295 MW; 1C5CD4013CB365E CRC64;

```

```

O97518
ID O97518 PRELIMINARY; PRT; 95 AA.
AC O97518;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deindard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010299; AAC67227.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . ; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
FT NON TER
SQ SEQUENCE 95 AA; 9221 MW; 96FB63A12B8A3647 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88
|||||

O97518
ID O97518 PRELIMINARY; PRT; 99 AA.
AC O97518;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deindard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010298; AAC67226.1; -.

```



```
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0001786; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCRHOPOPSN.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
SQ SEQUENCE 99 AA; 9535 MW; D4CED9750184BC47 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 99;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 85 LLLALLVL 92

RESULT 38
QBE9P1
ID QBE9P1 PRELIMINARY; PRT; 104 AA.
AC QBE9P1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cell division protein FtsL.
DE FtsL OR SO4226.
GN Shewanella oneidensis.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealeon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RL Nat. Biotechnol. 20:1118-1123 (2002).
DR EMBL; AS015855; AAN57198.1; -.
DR TIGR; SO4226; -.
DR InterPro: IPR007082; FtsL.
DR Pfam; PF04999; FtsL; 1.
KW Complete proteome.
SQ SEQUENCE 104 AA; 12187 MW; 408B210954032920 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 104;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 23 LLLALLVL 30

RESULT 39
Q9SDS4
ID Q9SDS4 PRELIMINARY; PRT; 106 AA.
AC Q9SDS4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Non-specific lipid transfer protein.

GN GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0001786; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCRHOPOPSN.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
SQ SEQUENCE 99 AA; 9535 MW; D4CED9750184BC47 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 99;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 85 LLLALLVL 92

RESULT 39
Q9SDS4
ID Q9SDS4 PRELIMINARY; PRT; 112 AA.
AC Q9SDS4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Salivary androgen-binding protein gamma
DE subunit).

GN LTP1.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamids; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Hanbyul; TISSUE=Leaf;
RA Hwang B.K., Jung H.W.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF208832; AAF23458.1; -.
DR HSSP; P23096; 1RZL.
DR GO: GO:0008289; P:lipid binding; IEA.
DR GO: GO:0006859; P:lipid transport; IEA.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR000528; Plant LTP.
DR Pfam; PF00234; tryptophan alpha amy1; 1.
DR PRINTS: PR00382; LIPIDTRANSF.
DR SMART; SM00499; AAI; 1.
DR SEQUENCE 106 AA; 11291 MW; 59C43238B9381C9 CRC64;

Query Match 0.7%; Score 8; DB 10; Length 106;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 GOIQGVV 204
DB 20 GOIQGVV 27

RESULT 40
Q8R1E9
ID Q8R1E9 PRELIMINARY; PRT; 112 AA.
AC Q8R1E9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN C2A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Salivary gland;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024677; AAH24677.1; -.
DR MGD; MGI:2655426; C2a.
DR InterPro: IPR006038; Uteroglobin_supf.
KW Hypothetical protein.
SQ SEQUENCE 112 AA; 12371 MW; 69CE957825183F72 CRC64;

Query Match 0.7%; Score 8; DB 11; Length 112;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
DB 5 LLLALLLV 12

RESULT 41
Q8JZX1
ID Q8JZX1 PRELIMINARY; PRT; 112 AA.
AC Q8JZX1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Salivary androgen-binding protein gamma
DE subunit).
```

C2D OR ABPG.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Salivary gland;  
RA Strausberg R.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C3H/HeJ;  
RX MEDLINE=22680631; PubMed=12795612;  
RA Karn R.C., Laukaitis C.M.;  
RT "Characterization of two forms of mouse salivary androgen-binding  
protein (ABP): implications for evolutionary relationships and ligand-  
binding function.";  
RL Biochemistry 42:7162-7170(2003).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C3H/HeJ;  
RA Laukaitis C.M., Dlouhy S.R., Karn R.C.;  
RT "The mouse salivary androgen-binding protein (ABP) gene cluster on  
Chromosome 7: Characterization and evolutionary relationships.";  
RL Mamm. Genome 0:0-0(2003).  
DR EMBL; BC035222; AAP35222.1; -.  
DR EMBL; AY325898; AAP94990.1; -.  
DR MGD; MGI:2655745; C2d.  
KW Hypothetical protein.  
SQ SEQUENCE 112 AA; 12794 MW; F832B5868F4BD53D CRC64;  
Query Match 0.7%; Score 8; DB 11; Length 112;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1153 LLLLLLV 1160  
DB 5 LLLLLLV 12

RESULT 42  
ID Q7TNV5 PRELIMINARY; PRT; 112 AA.  
AC Q7TNV5;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Salivary androgen-binding protein beta subunit.  
GN ABPG.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C3H/HeJ;  
RX MEDLINE=22680631; PubMed=12795612;  
RA Karn R.C., Laukaitis C.M.;  
RT "Characterization of two forms of mouse salivary androgen-binding  
protein (ABP): implications for evolutionary relationships and ligand-  
binding function.";  
RL Biochemistry 42:7162-7170(2003).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C3H/HeJ;  
RA Laukaitis C.M., Dlouhy S.R., Karn R.C.;  
RT "The mouse salivary androgen-binding protein (ABP) gene cluster on  
Chromosome 7: Characterization and evolutionary relationships.";  
RL Mamm. Genome 0:0-0(2003).  
DR EMBL; AY325897; AAP94989.1; -.  
SQ SEQUENCE 112 AA; 12383 MW; EB5F2C68DBF8D17C CRC64;

Query Match 0.7%; Score 8; DB 11; Length 112;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1153 LLLLLLV 1160  
DB 5 LLLLLLV 12

RESULT 43  
ID Q7TNN0 PRELIMINARY; PRT; 112 AA.  
AC Q7TNN0;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Salivary androgen-binding protein beta subunit.  
GN ABPG.  
OS Mus musculus (eastern European house mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=39442;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DBA/2J;  
RX MEDLINE=22680631; PubMed=12795612;  
RA Karn R.C., Laukaitis C.M.;  
RT "Characterization of two forms of mouse salivary androgen-binding  
protein (ABP): implications for evolutionary relationships and ligand-  
binding function.";  
RL Biochemistry 42:7162-7170(2003).  
DR EMBL; AY293278; AAP44465.1; -.  
SQ SEQUENCE 112 AA; 12411 MW; 818431A8CF8C029 CRC64;  
Query Match 0.7%; Score 8; DB 11; Length 112;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1153 LLLLLLV 1160  
DB 5 LLLLLLV 12

RESULT 44  
ID Q7TNW7 PRELIMINARY; PRT; 112 AA.  
AC Q7TNW7;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Salivary androgen-binding protein gamma subunit.  
GN ABPG.  
OS Mus musculus (eastern European house mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=39442;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DBA/2J;  
RX MEDLINE=22680631; PubMed=12795612;  
RA Karn R.C., Laukaitis C.M.;  
RT "Characterization of two forms of mouse salivary androgen-binding  
protein (ABP): implications for evolutionary relationships and ligand-  
binding function.";  
RL Biochemistry 42:7162-7170(2003).  
DR EMBL; AY293281; AAP44468.1; -.  
SQ SEQUENCE 112 AA; 12546 MW; 9F955317D9116883 CRC64;  
Query Match 0.7%; Score 8; DB 11; Length 112;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1153 LLLLLLV 1160

```

DR TIGR; DR2210; -.
KW Hypothetical protein; Complete proteome
SQ SEQUENCE 127 AA; 14314 MW; 2B3E79D3522A081 CRC64;

Query Match      0.7%; Score 8; DB 16; Length 127;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159
Db 7 GLLLLALL 14

RESULT 47
Q82U09 PRELIMINARY; PRT; 127 AA.
AC Q82U09
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 25, Last annotation update)
DE CRCB protein.
GN NEI704.
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19718 / IF0 14298;
RX MEDLINE=22586410; PubMed=12700255;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA Arciero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;
RA "Complete genome sequence of the ammonia-oxidizing bacterium and
RT obligate chemolithoautotroph Nitrosomonas europaea.";
RL J. Bacteriol. 185:2759-2773(2003).
DR EMBL; BX321862; CAD85615.1; -.
DR GO; GO:0016020; C-membrane; IEA.
DR InterPro; IPR003691; Camphor_CrbB.
DR Pfam; PF02537; CRCB; 1.
DR TIGRPFAMs; TIGR00494; crCB; 1.
KW Complete proteome.
SQ SEQUENCE 127 AA; 13419 MW; 55F30C8DD8AD7952 CRC64;

Query Match      0.7%; Score 8; DB 16; Length 127;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1147 GSTLGGLL 1154
Db 11 GSTLGGLL 18

RESULT 48
Q8C442 PRELIMINARY; PRT; 137 AA.
AC Q8C442
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Stimulated by retinoic acid gene 6.
GN STRA6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Hippocampus;
RX MEDLINE=22354683; PubMed=12466851;
RA THE FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of

```

```

Db 5 LLLALLV 12

RESULT 45
Q9ZPI9 PRELIMINARY; PRT; 114 AA.
AC Q9ZPI9
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Lipid transfer protein LTP1.
GN LTP1.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. VK-1;
RA Park C.J., Shin R., Paek K.H.;
RT "Characterization of a lipid transfer protein (CaiLTP1) in pepper.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF118131; AAD18029.1; -.
DR HSSP; P23036; 1RZL.
DR GO; GO:0008289; P:lipid binding; IEA.
DR GO; GO:0008689; P:lipid transport; IEA.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; tryP_alpha_1; 1.
DR PRINTS; PR00382; LIPIDTRANSF.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 114 AA; 12114 MW; 9315BA2125A83012 CRC64;

Query Match      0.7%; Score 8; DB 10; Length 114;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 197 GOIQGVV 204
Db 28 GOIQGVV 35

RESULT 46
Q9RSB5 PRELIMINARY; PRT; 127 AA.
AC Q9RSB5
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein DR2210.
GN DR2210.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13919 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey B.K., Peterson J.D.,
RA Dodson K.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vanathavan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE002054; AAF11763.1; -.
DR PIR; B75301; B75301.

```

```
RT 60,770 full-length cDNAs.;
RL Nature 420:563-573(2002).;
DR EMBL; AK083120; BAC38769.1; -.
DR PIR; PT0714; PT0714.
DR MGD; MGI:107742; Strag.
SQ SEQUENCE 137 AA; 14573 MW; 27A54C974EF867CC CRC64;

Query Match 0.7%; Score 8; DB 11; Length 137;
Best Local Similarity 100.0%; Pred.No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160
Db 64 LLLLLLV 71

RESULT 49
Q61692 PRELIMINARY; PRT; 141 AA.
AC Q61692;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative heat stable antigen.
GN HSA-C.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBA X C57BL/6; TISSUE=Spleen;
RX MEDLINE=91209380; Pubmed=2019286;
RA Wenger R.H., Avane M., Bose R., Koehler G., Nielsen P.J.;
RT "The genes for a mouse hematopoietic differentiation marker called the
RT heat-stable antigen.";
RL Eur. J. Immunol. 21:1039-1046(1991).
DR EMBL; X55486; CAA39843.1; -.
DR PIR; S15785; S15785.
SQ SEQUENCE 141 AA; 15515 MW; B4BFB428ADC03C69 CRC64;

Query Match 0.7%; Score 8; DB 11; Length 141;
Best Local Similarity 100.0%; Pred.No. 30;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 12 GLLLLALL 19

RESULT 50
Q7T1M3 PRELIMINARY; PRT; 144 AA.
AC Q7T1M3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bradykinin-potentiating/C-type natriuretic peptide isoform 2
DE (Fragment).
OS Bothrops jararacussu (Jararacussu).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
CX NCBI_TaxID=8726;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Hayashi M.A.F., Lameu C., Radis-Baptista G., Yamane T.,
RA Camargo A.C.M.;
RT "Cloning and sequence analysis of a Bothrops jararacussu BPPS
RT precursor.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY310915; AAP83422.1; -.

FT NON TER 144 144
SQ SEQUENCE 144 AA; 15982 MW; B23FED5230A84910 CRC64;

Query Match 0.7%; Score 8; DB 13; Length 144;
Best Local Similarity 100.0%; Pred.No. 30;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 10 GLLLLALL 17

RESULT 51
Q8QG90 PRELIMINARY; PRT; 146 AA.
AC Q8QG90;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Bradykinin-potentiating protein.
OS Bothrops insularis (Island jararaca) (Queimada jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
CX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Junqueira-de-Azevedo I.L.M., Ho P.L.;
RT "A survey of gene expression and diversity in the venom glands of the
RT pitviper Bothrops insularis through the generation of Expressed
RT Sequence Tags (ESTs).";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF490532; AAM09691.1; -.
SQ SEQUENCE 146 AA; 16088 MW; F58244C4F24B1799 CRC64;

Query Match 0.7%; Score 8; DB 13; Length 146;
Best Local Similarity 100.0%; Pred.No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 10 GLLLLALL 17

RESULT 52
Q7XK40 PRELIMINARY; PRT; 161 AA.
AC Q7XK40;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE OSUNBA0044X18.5 protein.
GN OSUNBA0044X18.5.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
CX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,
RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,
RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,
RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu P.H., Han B., Feng Q.,
RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Wu J., Yu Z.,
RA Chen L., Fan D.L., Wang Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,
RA Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,
RA Qian Y.M., Yifeng K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,
RA Zhang R.Q., Guan J.P., Hong G.F.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL606595; CAE05863.1; -.
SQ SEQUENCE 161 AA; 16300 MW; 0B8B2DF9C58ADC61 CRC64;
```

```
Query Match      0.7%; Score 8; DB 10; Length 161;
Best Local Similarity 100.0%; Pred.No.33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1156 LALLVLLAL 1163
DB 136 LALLVLLAL 143

RESULT 53
Q98MM8 PRELIMINARY; PRT; 174 AA.
ID Q98MM8
AC Q98MM8
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein mlr0512.
GN MLR0512.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Katanabe A., Ideawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP002995; BAB48085.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 174 AA; 20270 MW; E91E12118C11022F CRC64;

Query Match      0.7%; Score 8; DB 16; Length 174;
Best Local Similarity 100.0%; Pred.No.36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 LRLSLKALK 1097
DB 38 LRLSLKALK 45

RESULT 54
P79786 PRELIMINARY; PRT; 181 AA.
ID P79786
AC P79786
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Thrombospondin-1 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
EX MEDLINE=97209337; PubMed=9056637;
RA Tucker R.P., Hagios C., Chiquet-Ehrismann R., Lawler J.;
RT "In situ localization of thrombospondin-1 and thrombospondin-3
RT transcripts in the avian embryo.";
RL Dev. Dyn. 208:326-337(1997).
DR EMBL; U76994; AAB19208.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003367; tsep_3.
DR Pfam; PF02412; tsep_3; 8.

Query Match      0.7%; Score 8; DB 13; Length 181;
Best Local Similarity 100.0%; Pred.No.37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLALL 1159
DB 10 GLLLALL 17

RESULT 56
Q90Y11 PRELIMINARY; PRT; 181 AA.
ID Q90Y11
AC Q90Y11
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Bradykinin potentiating peptide and C-type natriuretic peptide isoform
DE 2.
OS Crotalus durissus terrificus (South American rattlesnake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Crotalus.
OX NCBI_TaxID=8732;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Hayashi M.A.F., Radis-Baptista G., Barbosa S.R., Yamane T.,
RA Camargo A.C.M.;
RT "Crotales durissus terrificus bradykinin potentiating peptide
RT precursor.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF308593; AAL09426.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR000663; Natr_Peptide.
DR Pfam; PF00212; ANP; 1.
DR PRINTS; PR00710; NATPEPTIDES.
DR SMART; SM00183; NAT_PEP; 1.
DR PROSITE; PS00263; NATRIURETIC_PEPTIDE; 1.
SQ SEQUENCE 181 AA; 18560 MW; 7B5ADCSB9372D07F CRC64;

Query Match      0.7%; Score 8; DB 13; Length 181;
Best Local Similarity 100.0%; Pred.No.37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLALL 1159
DB 10 GLLLALL 17

RESULT 56
Q90Y11 PRELIMINARY; PRT; 181 AA.
ID Q90Y11
AC Q90Y11
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Bradykinin potentiating peptide and C-type natriuretic peptide isoform
DE 2.
OS Crotalus durissus terrificus (South American rattlesnake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Crotalus.
OX NCBI_TaxID=8732;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Hayashi M.A.F., Radis-Baptista G., Barbosa S.R., Yamane T.,
RA Camargo A.C.M.;
```

RT \*Crotalus durissus terrificus bradykinin-potentiating peptide and C-  
RT type natriuretic peptide precursor isoform2.\*;  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF308594; AAL09427.1; -;  
DR GO: GO:0005576; C:extracellular; IEA.  
DR GO: GO:0005179; F:hormone activity; IEA.  
DR InterPro: IPR000663; Natr\_peptide.  
DR Pfam: PF00212; ANP; 1.  
DR PRINTS: PRO0710; NATPEPTIDES.  
DR SMART: SM00183; NAT\_PEP; 1.  
DR PROSITE: PS00263; NATRIURETIC\_PEPTIDE; 1.  
SQ SEQUENCE 181 AA; 18507 MW; 9B2B95D38AA5FF27 CRC64;

Query Match 0.7%; Score 8; DB 13; Length 181;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159  
DB 10 GLLLLALL 17

RESULT 57

Q8U4N4 PRELIMINARY; PRT; 222 AA.  
AC Q8U4N4;  
DT 01-JUN-2002 (T-REMBLrel. 21, Created)  
DT 01-JUN-2002 (T-REMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (T-REMBLrel. 21, Last annotation update)  
DE Hypothetical protein PF0047.  
GN PF0047.  
OS Pyrococcus furiosus.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Pyrococcus.  
OX NCBI\_TaxID=2261;  
RN [1]  
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;  
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;  
RT "The complete sequence of the Pyrococcus furiosus genome.";  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AE010131; AAL80171.1; -;  
RW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 222 AA; 24556 MW; 721D79ED547A0F13 CRC64;

Query Match 0.7%; Score 9; DB 17; Length 222;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 512 RGRVYVE 519  
DB 106 RGRVYVE 113

RESULT 58

Q7V5L4 PRELIMINARY; PRT; 224 AA.  
AC Q7V5L4;  
DT 01-OCT-2003 (T-REMBLrel. 25, Created)  
DT 01-OCT-2003 (T-REMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (T-REMBLrel. 25, Last annotation update)  
DE Conserved hypothetical membrane protein.  
GN P3H9.20 OR PWT1539.  
OS Prochlorococcus marinus (strain MIT 9313).  
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
OC Prochlorococcus.  
OX NCBI\_TaxID=74547;  
RN [1]  
RC SEQUENCE FROM N.A.  
RX MEDLINE=22825698; PubMed=12917642;  
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,  
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,  
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,

RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,  
RA Webb E.A., Zinser E.R., Chisholm S.W.;  
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic  
RT niche differentiation.";  
RL Nature 424:1042-1047(2003).  
DR EMBL: BX572099; CA321714.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 224 AA; 24172 MW; E1CD9F149C279A29 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 224;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1158 LVLALRK 1165  
DB 190 LVLALRK 197

RESULT 59

OS7835 PRELIMINARY; PRT; 224 AA.  
AC OS7835;  
DT 01-AUG-1998 (T-REMBLrel. 07, Created)  
DT 01-AUG-1998 (T-REMBLrel. 07, Last sequence update)  
DT 01-JUN-2003 (T-REMBLrel. 24, Last annotation update)  
DE Hypothetical protein PH0087.  
GN PH0087.  
OS Pyrococcus horikoshii.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Pyrococcus.  
OX NCBI\_TaxID=53953;  
RN [1]  
RC SEQUENCE FROM N.A.  
RX STRAIN=OT3;  
RY MEDLINE=98344137; PubMed=9679194;  
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,  
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
RA Masuchi Y., Shizuya H., Kikuchi H.;  
RT "Complete sequence and gene organization of the genome of a hyper-  
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA Res. 5:55-76(1998).  
DR EMBL: AP000001; BAA29156.1; -;  
DR PIR: E71228; E71228.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 224 AA; 25134 MW; 9CC56212064558E0 CRC64;

Query Match 0.7%; Score 8; DB 17; Length 224;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 512 RGRVYVE 519  
DB 105 RGRVYVE 112

RESULT 60

P74604 PRELIMINARY; PRT; 226 AA.  
ID P74604;  
AC P74604;  
DT 01-FEB-1997 (T-REMBLrel. 02, Created)  
DT 01-FEB-1997 (T-REMBLrel. 02, Last sequence update)  
DT 01-JUN-2003 (T-REMBLrel. 24, Last annotation update)  
DE Hypothetical protein slr1571.  
GN SLR1571.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RC SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;

RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirosewa M., Sugiyama M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.,  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synecocystis sp. strain PC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.",  
RL DNA Res. 3:109-136(1996).  
DR EMBL; D90916; BAA18712.1; -.  
DR PIR; S76800; S76800.  
DR InterPro; IPR002838; DUF124.  
DR Pfam; PF01987; DUF124; 1.  
DR ProDom; PD013634; DUF124; 1.  
DR TIGRFam; TIGR00266; TIGR00266; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 226 AA; 25003 MW; 29C5C137C128DAC0 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 226;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1073 RLVPNQEI 1080  
Db 183 RLVPNQEI 190  
|||||

RESULT 61  
Q882P2 PRELIMINARY; PRT; 227 AA.  
AC Q882P2;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Conserved hypothetical protein.  
GN PSPT02584.  
OS Pseudomonas syringae (pv. tomato).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=323;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DC3000;  
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,  
RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,  
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,  
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidson T.,  
RA White O., Fraser C., Collier A.;  
RT "Complete sequence of Pseudomonas syringae."  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF016865; AAC056088.1; -.  
DR TIGR; PSPT02584; -.  
DR InterPro; IPR002489; DUF14.  
DR Pfam; PF01493; GXKG; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 227 AA; 24036 MW; AB3458F347BEP145 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 227;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 547 ASVRDLNQ 554  
Db 9 ASVRDLNQ 16  
|||||

RESULT 62  
Q9AJX4 PRELIMINARY; PRT; 231 AA.  
ID Q9AJX4  
AC Q9AJX4;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Putative oxidoreductase.  
GN SC01803 OR SCI33.02.  
OS Streptomyces coelicolor.  
OC Bacteria; Actinobacteria; Actinomycetales; Actinomycetales;  
OC Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Harris D.;  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RX MEDLINE=97000351; PubMed=8843436;  
RA Redenbach M., Kieser H.M., Denapait D., Eichner A., Cullum J.,  
RA Kinashi H., Hopwood D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL Mol. Microbiol. 21:77-96(1996).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2) / M145;  
RX MEDLINE=21996410; PubMed=12000953;  
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
RA Thomson N.R., James K.P., Harris D.E., Quail M.A., Kieser H.,  
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
RA Hopwood D.A.;  
RT "Complete genome sequence of the model actinomycete Streptomyces  
RT coelicolor A3(2).";  
RL Nature 417:141-147(2002).  
CC - SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
CC (SDR) FAMILY.  
DR EMBL; AL939110; CAC28543.1; -.  
DR HSP; O64105; 10AA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR002198; ADH\_short.  
DR Pfam; PF00106; adh\_short; 1.  
DR PRINTS; PR00080; SDRFAMILY.  
DR PROSITE; PS00061; ADH\_SHORT; 1.  
KW Oxidoreductase; Complete proteome.  
SQ SEQUENCE 231 AA; 23789 MW; 6AEDB40306B94085 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 231;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 790 DLVLDRS 797  
Db 27 DLVLDRS 34  
|||||

RESULT 63  
Q98GF8 PRELIMINARY; PRT; 235 AA.  
ID Q98GF8  
AC Q98GF8;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE ATP-binding component of phosphate transport, HisP-like  
DE nucleotide binding protein, PhnL.  
GN MLR3349.

OS Rhizobium loti (Mesorhizobium loti).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_TaxID=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAFF303099;  
 RX MEDLINE=21082930; PubMed=11214969;  
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
 RA Watanabe A., Igesawa K., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
 RA Moichizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 Mesorhizobium loti";  
 RL DNA Res. 7:331-338(2000).  
 DR EMBL; AF003001; BAB50258.1; -;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR003439; ABC\_TRANSPORTER.  
 DR Pfam; PF00005; ABC\_tran; 1.  
 DR ProDom; PD00006; ABC\_transporter; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 1.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 1.  
 KW ATP-binding; Complete proteome.  
 SQ SEQUENCE 235 AA; 25098 MW; 6FA8F03F3F827361 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 235;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 EDGLIDLA 623  
 DB 73 EDGLIDLA 80

RESULT 64  
 Q02020 PRELIMINARY; PRT; 245 AA.  
 AC Q02020;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein F40F4.7.  
 GN F40F4.7.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pterodermidae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA Waterston R.;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Wilson R.;  
 RT "The sequence of C. elegans cosmid F40F4.";  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; U00420; AAK84526.3; -;  
 DR WormPep; F40F4.7; CE28316.  
 DR GO; GO:0008080; F:N-acetyltransferase activity; IEA.  
 DR InterPro; IPR00182; GCM5acetyl\_trans.  
 DR Pfam; PF00583; Acetyltrans; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 245 AA; 27893 MW; AS3289FBB2F91E75 CRC64;

Query Match 0.7%; Score 8; DB 5; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 557 YNDVVVGA 564  
 DB 141 YNDVVVGA 148

RESULT 65  
 Q8Y1R9 PRELIMINARY; PRT; 257 AA.  
 AC Q8Y1R9;  
 DT 01-VAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-VAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein RSC0620.  
 GN RSC0620 OR RS01521.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Burkholderiaceae; Ralstonia.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GMI1000;  
 RX MEDLINE=21681879; PubMed=11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,  
 RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,  
 RA Signier P., Thebault P., Whalen M., Wincker P., Levy M.,  
 RA Weissenbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";  
 RL Nature 415:497-502(2002).  
 DR EMBL; AL646060; CAD14150.1; -;  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000572; Oxidored\_molyp.  
 DR Pfam; PF00174; oxidored\_molyp; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 257 AA; 29260 MW; DAE569EE1572113D CRC64;

Query Match 0.7%; Score 8; DB 16; Length 257;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1149 TLGGLLLL 1156  
 DB 37 TLGGLLLL 44

RESULT 66  
 Q8Q991 PRELIMINARY; PRT; 265 AA.  
 AC Q8Q991;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Bradykinin-potentiating/c-type natriuretic protein.  
 OS Bothrops insularis (Island jararaca) (Queimada jararaca).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidosauria; Squamata; Serpentes; Colubroidea;  
 OC Viperidae; Crotalinae; Bothrops.  
 OX NCBI\_TaxID=8723;  
 RN [1]  
 RP SEQUENCE FROM N.A.



```

RC TISSUE=Venom gland;
RA Junqueira-de-Azevedo I.L.M., Ho P.L.;
RT "A survey of gene expression and diversity in the venom glands of the
RT pitviper Bothrops insularis through the generation of Expressed
RT Sequence Tags (ESTs)."; to the EMBL/GenBank/DBJ databases.
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF490531; AAM09690.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR000663; Natr_peptide.
DR Pfam; PF00212; ANP; 1.
DR PRINTS; PR00710; NATPEPTIDES.
DR SMART; SM00183; NAT_PEP; 1.
DR PROSITE; PS00263; NATRIURETIC_PEPTIDE; 1.
SQ SEQUENCE 265 AA; 27763 MW; 0AE1408B42358BE CRC64;

Query Match 0.7%; Score 8; DB 13; Length 265;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 10 GLLLLALL 17

RESULT 67
Q9PW56 PRELIMINARY; PRT; 265 AA.
AC Q9PW56;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Bradykinin-potentiating/C-type natriuretic peptide.
OS Bothrops jararaca (Jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8724;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Hayashi M.A.F., Murbach A.F., Camargo A.C.M.;
RT "The precursor of C-type natriuretic peptide of snake brain contains
RT angiotensin converting enzyme inhibitors, specific for the C-catalytic
RT site."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF171670; AAD51326.2; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR000663; Natr_peptide.
DR Pfam; PF00212; ANP; 1.
DR PRINTS; PR00710; NATPEPTIDES.
DR SMART; SM00183; NAT_PEP; 1.
DR PROSITE; PS00263; NATRIURETIC_PEPTIDE; 1.
SQ SEQUENCE 265 AA; 27763 MW; 8E99AEC976CCD439 CRC64;

Query Match 0.7%; Score 8; DB 13; Length 265;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 10 GLLLLALL 17

RESULT 68
Q861M5 PRELIMINARY; PRT; 272 AA.
AC Q861M5;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE MHC class I antigen (Fragment).

```

```

OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=107;
RA Chung C., Leib S.R., Fraser D.G., Ellis S.A., McGuire T.C.;
RT "Sequence-based typing of horse classical MHC class I genes."
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY176095; AAO18112.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006955; F:immune response; IEA.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG_c1.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR SMART; SM00407; IGc1; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
FT NON_TER 1
FT NON_TER 272
SQ SEQUENCE 272 AA; 30576 MW; 8EC9B0C183D319FD CRC64;

Query Match 0.7%; Score 8; DB 7; Length 272;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 VLLGAVGA 387
DB 232 VLLGAVGA 239

RESULT 69
Q861K3 PRELIMINARY; PRT; 272 AA.
AC Q861K3;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE MHC class I antigen (Fragment).
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=130;
RA Chung C., Leib S.R., Fraser D.G., Ellis S.A., McGuire T.C.;
RT "Sequence-based typing of horse classical MHC class I genes."
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY176118; AAO18134.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006955; F:immune response; IEA.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG_c1.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR PRODOM; PD000050; MHC_I; 1.
DR SMART; SM00407; IGc1; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
FT NON_TER 1
FT NON_TER 272
SQ SEQUENCE 272 AA; 30220 MW; F7FC66FFB06558CF CRC64;

Query Match 0.7%; Score 8; DB 7; Length 272;

```

```
Best Local Similarity 100.0%; Pred. No. 53;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 VLLGAVCA 387
Db 232 VLLGAVCA 239

RESULT 70
Q8P8NG PRELIMINARY; PRT; 278 AA.
AC Q8P8NG
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE ABC transporter sugar permease.
GN LAGC OR KCC204.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Terro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannaval F., Cardoso J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.P., Takita M.A., Tamura R.E., Teixeira B.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White P.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AE012328; AM41484.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD_transp.
DR Pfam; PF00528; BPD_transp; 1.
RW Complete proteome.
SQ SEQUENCE 278 AA; 30673 MW; FB45244D924B7361 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 278;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLAL 1158
Db 17 GGLLLAL 24

RESULT 71
Q8TPV4 PRELIMINARY; PRT; 290 AA.
AC Q8TPV4
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Uncharacterized domain specific for M.kandleri, MK-34 family, a.
GN MK1344.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=381;
```

```
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezheva V.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozlovskiy S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL; AE010427; AA02557.1; -.
KW Complete proteome.
SQ SEQUENCE 290 AA; 32754 MW; 706E6883976ADF91 CRC64;

Query Match 0.7%; Score 8; DB 17; Length 290;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 268 VITDGESH 275
Db 161 VITDGESH 168

RESULT 72
Q94K03 PRELIMINARY; PRT; 299 AA.
AC Q94K03;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Unknown protein.
GN A IG002N01.8
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF370497; AA43874.1; -.
SQ SEQUENCE 299 AA; 34028 MW; BB2AB03C4A184D3E CRC64;

Query Match 0.7%; Score 8; DB 10; Length 299;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
Db 44 LLLALLLV 51

RESULT 73
Q98BD8 PRELIMINARY; PRT; 307 AA.
AC Q98BD8;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein ml15620.
GN ML15620.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
```

```

RN  SEQUENCE FROM N.A.
RP  STRAIN=VAFF303099;
RX  MEDLINE=21082930; PubMed=11214968;
RA  Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA  Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA  Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA  Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA  Takeuchi C., Yamada M., Tabata S.;
RT  "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT  Mesorhizobium loti.";
RL  DNA Res. 7:331-338(2000).
DR  EMBL; AP003007; BAB52034.1; -.
DR  InterPro; IPR002781; DUF81.
DR  Pfam; PF01925; DUF81; 1.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 307 AA; 31951 MW; 938FFC1D38C4AC16 CRC64;

Query Match      0.7%; Score 8; DB 16; Length 307;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1155 LLALLVLA 1162
DB  275 LLALLVLA 282
|||||

RESULT 74
Q92885 PRELIMINARY; PRT; 308 AA.
AC  Q92885;
DT  01-DEC-2001 (TrEMBLrel. 19, Created)
DT  01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  Hypothetical transmembrane protein SMC02232.
GN  R00337 OR SMC02232.
OS  Rhizobium meliloti (Sinorhizobium meliloti).
OC  Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC  Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX  NCBI_TaxID=382;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=1021;
RX  MEDLINE=21396507; PubMed=11481430;
RA  Capela D., Barloy-Hubier F., Guzy J., Bothe G., Ampe F., Batut J.,
RA  Boistard P., Becker A., Boutery M., Cadieu E., Dreano S., Gloux S.,
RA  Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA  Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA  Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT  "Analysis of the chromosome sequence of the legume symbiont
RT  Sinorhizobium meliloti strain 1021.";
RL  Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR  EMBL; AL591784; CAC45109.1; -.
DR  InterPro; IPR002781; DUF81.
DR  Pfam; PF01925; DUF81; 1.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 308 AA; 32465 MW; C8497E6E16BC0D CRC64;

Query Match      0.7%; Score 8; DB 16; Length 308;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1155 LLALLVLA 1162
DB  275 LLALLVLA 282
|||||

RESULT 75
Q8CB84 PRELIMINARY; PRT; 323 AA.
AC  Q8CB84;
DT  01-MAR-2003 (TrEMBLrel. 23, Created)
DT  01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

```

```

DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  Integrin alpha 2 (Fragment).
GN  ITGA2.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Bone;
RX  MEDLINE=22354683; PubMed=12466851;
RA  The FANTOM Consortium,
RA  The RIKEN Genome Exploration Research Group Phase I & II Team;
RT  "Analysis of the mouse transcriptome based on functional annotation of
RT  60,770 full-length cDNAs.";
RL  Nature 420:563-573(2002).
DR  EMBL; AK036572; BAC29485.1; -.
DR  MGD; MGI:96600; Itga2.
DR  GO; GO:0008305; C:integrin complex; IEA.
DR  GO; GO:0004895; P:cell adhesion receptor activity; IEA.
DR  GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR  InterPro; IPR000413; Integrin_alpha.
DR  InterPro; IPR002035; VWF_A.
DR  Pfam; PF00092; vwa; 1.
DR  PRINTS; PR00453; VWFADOMAIN.
DR  SMART; SM00191; Int_alpha; 1.
DR  SMART; SM00327; VWA; 1.
DR  PROSITE; PS50234; VWFA; 1.
FT  NON_TER 323 323
SQ  SEQUENCE 323 AA; 35025 MW; 05A456A44283724A CRC64;

Query Match      0.7%; Score 8; DB 11; Length 323;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  71 GDVYKCPV 78
DB  75 GDVYKCPV 82
|||||

Search completed: June 24, 2004, 18:06:14
Job time : 63 secs

```



OM protein - protein search, using sw model

Run on: June 24, 2004, 17:56:51 ; Search time 65 Seconds  
(without alignments)  
5164.099 Million cell updates/sec

Title: US-09-980-403-2  
Perfect score: 1188  
Sequence: 1 MDLPRGLVAVALLSLWPGT.....FRSARRRPPGLDTPKVL 1188

Scoring table: OLIGO Gapop 60.0 , Gapext 60.0 1586107

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003Bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1188	100.0	1188	4	AAB30929 Amino aci
2	928	78.1	1189	4	ABG12949 Novel hum
3	926	77.9	1188	4	Aau14231 Human nov
4	926	77.9	1188	4	AAB50085 Human A25
5	926	77.9	1188	5	Aau10551 Human A25
6	911	76.7	1188	4	Aau14467 Human nov
7	911	76.7	1188	7	Ade09956 Novel pro
8	911	76.7	1189	3	AAB25582 ITGAl1 pr
9	911	76.7	1189	6	ABR58364 Human NOV
10	911	76.7	1189	6	ADA27054 Human nov
11	911	76.7	1189	7	ADG63570 Human pro
12	910	76.7	1189	8	ADG66584 Human nov
13	810	68.2	1034	3	AAB25590 Protein e
14	810	68.2	1034	6	ADA27062 Human nov
15	810	68.2	1034	8	ADG6592 Human hum
16	733	63.4	1120	6	ABR58365 Human NOV
17	489	41.2	707	4	Aau19663 Human nov
18	489	41.2	707	5	ABP47883 Human pol
19	489	41.2	707	7	ADG10845 Human ext
20	193	16.2	193	5	Aau176854 Human int
21	190	16.0	193	5	Aau176863 Human int
22	122	10.3	1188	4	AAB50087 Murine A2
23	122	10.3	1188	5	Aau10552 Murine A2
24	104	8.8	360	7	ADE08585 Novel pro
25	89	7.5	109	5	ABG66673 Human nov

Aaw65098 Canine be  
Aaw72832 Canine al

## ALIGNMENTS

## RESULT 1

AAB30929  
ID AAB30929 standard; protein: 1188 AA.

AC AAB30929;

DT 02-APR-2001 {first entry}

DE Amino acid sequence of a human  $\alpha 11$  integrin chain.

Human; integrin; alball subunit; fibroblast; muscle cell; chondrocyte;  
osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;  
wound healing; trauma; rheumatoid arthritis; osteoarthritis;  
osteoporosis; cartilage damage; bone damage; cartilage.

OS Homo sapiens.

FH	Key	Location/Qualifiers
----	-----	---------------------

FT Peptide 1. .22

ET		/note= "signal peptide"
ET		

FT region 951.972

FI	Domain	/NOTE=	1142	1164
FT				

```

FT      domain
FT      /note= "transmembrane domain"
FT      112..120

```

PN WO200075187-A1.

PD 14-DEC-2000.

PF 31-MAY-2000; 2000WO-SB001135.

PR 03-JUN-1999; 99SE-00002056.



PA (ACTI-) ACTIVE BIOTECH AB.

XX  
PT  
Cullbora D.

XX  
PT  
GUTHRIE D;

DR WPI: 2001-071061/08.

DR N-PSDB; AAC86871.

XX

PT Integrin subunit alpha 11 or integrin heterodimer comprising subunit alpha 11 in association with subunit beta, useful for treating muscle dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.

PS Disclosure; Fig 2a-c; 79pp; English.

The present sequence represents a human integrin subunit, designated alpha11. The alpha11 polynucleotide and polypeptide are useful as markers of cell target molecules, such as fibroblasts muscle cells, chondrocytes, osteoblasts, mesenchymally derived cells or stem cells. They are also used for determining the differential stage of cells during differentiation, development in pathological conditions, in tissue regeneration, in transplantation or in therapeutic and physiological repair of tissues. The pathological conditions involving subunit alpha11 are selected from damage of cells, muscle dystrophy, fibrosis, wound healing, trauma, rheumatoid arthritis, osteoarthritis and osteoporosis, damage of cartilage and bone, and cartilage and bone diseases. The polypeptide is useful for detecting the formation of cartilage during embryonic development, for detecting physiological therapeutic repair of cartilage and muscle, for selection and analysis, or for sorting, isolating or purification of chondrocytes and muscle cells, for detecting regeneration of cartilage or chondrocytes during transplantation of cartilage or chondrocytes during transplantation of cartilage or chondrocytes, respectively, or of muscle or muscle cells during transplantation of muscle or muscle cells, respectively, and for studies of differentiation of chondrocytes or muscle cells.

Db 961 RSSLSHVEVKLNSSLERYDGIQPPFSCIFRIQNLGLFPPIHGMKMTIPIATRSNRL 1020  
 QY 1021 KLRDLTDEANTSCNMGNSYRPTPVVEEDLRAPQNLHNSDVVSNCRNLRLVFNCEI 1080  
 Db 1021 KLRDLTDEANTSCNMGNSYRPTPVVEEDLRAPQNLHNSDVVSNCRNLRLVFNCEI 1080  
 QY 1081 NFHLGNLWLSLKALKYKSKIMVNAALQRFHSPPIFREDDPSRQIEFISKOEDWQV 1140  
 Db 1081 NFHLGNLWLSLKALKYKSKIMVNAALQRFHSPPIFREDDPSRQIEFISKOEDWQV 1140  
 QY 1141 PIWIVGSTLGLLLALLLALRLKLGFFRSARRRERGLDPTPKVLE 1188  
 Db 1141 PIWIVGSTLGLLLALLLALRLKLGFFRSARRRERGLDPTPKVLE 1188

## RESULT 2

ABGI12949  
 ID ABGI12949 standard; protein; 1189 AA.  
 XX AC ABGI12949;  
 XX 13-FEB-2002 (first entry)  
 DT Novel human diagnostic protein #12940.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX

OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS77136.

XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

PS Claim 20; SEQ ID NO 43308; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1189 AA;

Query Match 78.13; Score 928; DB 4; Length 1189;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1028; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDLPRGLVAVWALSLSLWPGFTDTFNMDTRKPRVPGSRTAFYGYTVQOEDISGNKWLWVGA 60  
 Db 1 MDLPRGLVAVWALSLSLWPGFTDTFNMDTRKPRVPGSRTAFYGYTVQOEDISGNKWLWVGA 60

QY 61 PLETNGYQKTDGVYKCPVIHGNCTKLNIGRVTLNSVSKRKNMRLGLSLATNPKNQNSFLA 120

Db 61 PLETNGYQKTDGVYKCPVIHGNCTKLNIGRVTLNSVSKRKNMRLGLSLATNPKNQNSFLA 120

QY 121 CSPLWSHECGSSVYTTGMCRSVNSNFRFSKTVAPALQRCQTYMDIVIVLDSGNSIYPWVE 180

Db 121 CSPLWSHECGSSVYTTGMCRSVNSNFRFSKTVAPALQRCQTYMDIVIVLDSGNSIYPWVE 180

QY 181 VQHFLINILKFFYIGPQIQGVVQVGEDVHFEHLNDYRSVKDVEAAASHIBORGTTET 240

Db 181 VQHFLINILKFFYIGPQIQGVVQVGEDVHFEHLNDYRSVKDVEAAASHIBORGTTET 240

QY 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSNHSPLDKVIQQSERDNNTRYAVAVL 300

Db 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSNHSPLDKVIQQSERDNNTRYAVAVL 300

QY 301 GYNRRGINPETFLNEIKYIASDDPKHFNVTDEAALKDIVDALGDRIFSLGNTNKNET 360

Db 301 GYNRRGINPETFLNEIKYIASDDPKHFNVTDEAALKDIVDALGDRIFSLGNTNKNET 360

QY 361 SFGLEMSOTGFSSHVVEDGVLGAVGAYDWNGAVLAKETSAGKVIPLRESYIKPEPELKN 420

Db 361 SFGLEMSOTGFSSHVVEDGVLGAVGAYDWNGAVLAKETSAGKVIPLRESYIKPEPELKN 420

QY 421 HGAYLGYTVTSVSSRQGRVYVAGAPRFNHTKGVLFTMNNRSLTIHQAMRGQIGSYF 480

Db 421 HGAYLGYTVTSVSSRQGRVYVAGAPRFNHTKGVLFTMNNRSLTIHQAMRGQIGSYF 480

QY 481 GSEITSVDDIDGCVTDVLLGAPMYNEGERGKVVYVELRQNFVYNGTLKQSHSYQNA 540

Db 481 GSEITSVDDIDGCVTDVLLGAPMYNEGERGKVVYVELRQNFVYNGTLKQSHSYQNA 540

QY 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGPGRGSLTKPKRITASELATG 600

Db 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGPGRGSLTKPKRITASELATG 600

QY 601 LQYFGCSIHQQLDNLNEDGLIDLAVGALGNVILWSRPVQINASLHFEPSKINIIFHRDCK 660

Db 601 LQYFGCSIHQQLDNLNEDGLIDLAVGALGNVILWSRPVQINASLHFEPSKINIIFHRDCK 660

QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYTPRAHLDGGDRFTNRVAV 720

Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYTPRAHLDGGDRFTNRVAV 720

QY 721 LLSSGQELCERINPHVLDADYKVPVTSVEYSLEDDPHGFMGLDDGPTTLRVSPFWNG 780

Db 721 LLSSGQELCERINPHVLDADYKVPVTSVEYSLEDDPHGFMGLDDGPTTLRVSPFWNG 780

QY 781 CNDEHCVPLDLVLDARSDDLTAKEYCORVLKPAQDCSAYTLGSDFTTVFIIESTRQAVAV 840

Db 781 CNDEHCVPLDLVLDARSDDLTAKEYCORVLKPAQDCSAYTLGSDFTTVFIIESTRQAVAV 840

QY 841 EATLENRGENAYSTVLNISQANLQFASLTQKSDSGSIECVNEERLQKQVCNVSPFF 900

Db 841 EATLENRGENAYSTVLNISQANLQFASLTQKSDSGSIECVNEERLQKQVCNVSPFF 900

QY 901 RAKAKVAFRLDSEFSKSIPLHLEIEILAAGSDSNERDSTKEDNVAPLRFHLKYEADVLF 960

Db 901 RAKAKVAFRLDSEFSKSIPLHLEIEILAAGSDSNERDSTKEDNVAPLRFHLKYEADVLF 960

QY 961 RSSLSHVEVKLNSSLERYDGIQPPFSCIFRIQNLGLFPPIHGMKMTIPIATRSNRL 1020

```

Db 961 RSSLSHYEVKLSLRYDGI GPPFSCIFRIQNLGLFPIHGMKTIPIATRSNRL 1020
Qy 1021 KLRDFTLDE 1029
Db 1021 KLRDFTLDE 1029

RESULT 3
AAU14231
ID AAU14231 standard; protein; 1188 AA.
AC AAU14231;
DT 24-OCT-2001 (first entry)
XX Human novel protein #102.
DE
KW Human; novel protein; Antianemic; osteopathic; antiinflammatory;
KW immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic;
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.
XX
OS Homo sapiens.
PN W0200155437-A2.
XX 02-AUG-2001.
PF 25-JAN-2001; 2001WO-US0002623.
XX 25-JAN-2000; 2000US-00491404.
PR (HYSE-) HYSEQ INC.
PA
PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-451939/48.
XX N-PSDB; AAS22536.
PT Isolated polypeptides useful for treating anti-inflammatory diseases,
PT nervous system disorders, and for regenerating bone and cartilage.
XX
XX Example 4; Page 578-581; 894pp; English.
XX
XX The invention relates to polynucleotides encoding novel human proteins or
XX their active domains. The polypeptides, polynucleotides and antibodies
XX raised against the polypeptides are used in a method of treatment of a
XX mammal and prevention of disorders caused by the aberrant protein
XX expression or activity. The polypeptides can be used as molecular weight
XX markers, food supplements, and in antibody production. The polypeptides
XX are used to identify compounds which bind to the polypeptides.
XX Polynucleotides of the invention are used as probes and primers, for
XX sequencing, for chromosome or gene mapping, in the production of
XX recombinant proteins, and in generating anti-sense DNA or RNA and in gene
XX therapy. Polypeptides of the invention can be used to target drugs to a
XX tumour, in assays to determine biological activity, to raise
XX antibodies/ elicit an immune response, to determine quantitative protein
XX levels, as tissue markers, and to isolate receptors or ligands.
XX Polypeptides of the invention may also be useful in treating platelet
XX disorders, stem cell disorders, regenerating bone, cartilage, tendon,
XX ligament and/or nerve tissue, wound healing, treating burns, promoting
XX the proliferation, differentiation and survival of stem cells, as a
XX contraceptive, treating osteoporosis and osteoarthritis, anaemia,
XX Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
XX sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
XX fungal infection or from autoimmunity, cancer, allergy, asthma, graft-
XX versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory
XX diseases, nervous system disorders, and infection. The present sequence
XX represents a protein of the invention.

```

```

XX Sequence 1188 AA;
SQ
Query Match 77.9%; Score 926; DB 4; Length 1188;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1126; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MDLPGLVAVWALSIPGFTDTFNMDTRKPRVIPSRTAFPGYTVQOHDISGNKLVVGA 60
Db 1 MDLPGLVAVWALSIPGFTDTFNMDTRKPRVIPSRTAFPGYTVQOHDISGNKLVVGA 60
Qy 61 PLETNGYQKTGDVYKCPVIHGNCTKLNIGRVTLSNVSRKDNMRGLGLSLATNPKDNSFLA 120
Db 61 PLETNGYQKTGDVYKCPVIHGNCTKLNIGRVTLSNVSRKDNMRGLGLSLATNPKDNSFLA 120
Qy 121 CSPLSHSCGSSYYTKGCSRVSNSFRSKTVAPALQRCQTTMDIIVLDGNSNYPWVE 180
Db 121 CSPLSHSCGSSYYTKGCSRVSNSFRSKTVAPALQRCQTTMDIIVLDGNSNYPWVE 180
Qy 181 VQHFLINILKXFYIGFGQIQGVVQYGDVVEHFLNDYRSVKQVVEAASHIEORGGTET 240
Db 181 VQHFLINILKXFYIGFGQIQGVVQYGDVVEHFLNDYRSVKQVVEAASHIEORGGTET 240
Qy 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITGESHSDPDLKYIOQSERDNVTRVAVL 300
Db 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITGESHSDPDLKYIOQSERDNVTRVAVL 300
Qy 301 GYNNRRGINPETFLNEIKYIASDPDDKHFFNVVTDAAALKDIVDALGDRIFSLGNTKNET 360
Db 301 GYNNRRGINPETFLNEIKYIASDPDDKHFFNVVTDAAALKDIVDALGDRIFSLGNTKNET 360
Qy 361 SFGLEMSQTFSSHVVEDGVLGAVGWNGAVLKETSAGKVIPLRSYLYKEPFEELKN 420
Db 361 SFGLEMSQTFSSHVVEDGVLGAVGWNGAVLKETSAGKVIPLRSYLYKEPFEELKN 420
Qy 421 HGAYLGYVTSVSSRQGRVVVAGAPRNHTGKVLFTPMHNRSLTIHQAMEGQOIGSYF 480
Db 421 HGAYLGYVTSVSSRQGRVVVAGAPRNHTGKVLFTPMHNRSLTIHQAMEGQOIGSYF 480
Qy 481 GSEITSDVDIGDGVTDVLLVGAPMYFNGRGRKVVYVELQNRVRYNGTLKDSHSYQNA 540
Db 481 GSEITSDVDIGDGVTDVLLVGAPMYFNGRGRKVVYVELQNRVRYNGTLKDSHSYQNA 540
Qy 541 RFGSSIASVRDLNDSYNDVVVVGAPLEDNHAGAIYIFHGFRGSIILTKPQRIATASELATG 600
Db 541 RFGSSIASVRDLNDSYNDVVVVGAPLEDNHAGAIYIFHGFRGSIILTKPQRIATASELATG 600
Qy 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILMSRPVQVQINASLHFPFSKINIFHRDCK 660
Db 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILMSRPVQVQINASLHFPFSKINIFHRDCK 660
Qy 661 RSGRDATCLAAFLCPTPIFLAPHFQTTTIGIRYNATMDERRVTPRAHLDGSDRFTNRAV 720
Db 661 RSGRDATCLAAFLCPTPIFLAPHFQTTTIGIRYNATMDERRVTPRAHLDGSDRFTNRAV 720
Qy 721 LLSQQELCERINFHVLDTADYVKPVTFSVEYSLEDPDHGMPLDDGWFTTLRSVPFWNG 780
Db 721 LLSQQELCERINFHVLDTADYVKPVTFSVEYSLEDPDHGMPLDDGWFTTLRSVPFWNG 780
Qy 781 CNEDEHCVFDVLVDARSOLPTAMEYQVRLKPAQDCSAYTILSPDTTFFIESTFORVAV 840
Db 781 CNEDEHCVFDVLVDARSOLPTAMEYQVRLKPAQDCSAYTILSPDTTFFIESTFORVAV 840
Qy 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIKEDNEERLRKQVCNVSYPFF 900
Db 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIKEDNEERLRKQVCNVSYPFF 900
Qy 901 RAKAKVAFRLDSERSKSIPLHLEIELAAGSDSNERDSTKEDNVAPAPHLKYEADVLT 960
Db 901 RAKAKVAFRLDSERSKSIPLHLEIELAAGSDSNERDSTKEDNVAPAPHLKYEADVLT 960
Qy 961 RSSLSHYEVKLSLRYDGI GPPFSCIFRIQNLGLFPIHGMKTIPIATRSNRL 1020

```



```

Db 961 RSSLSHYEVKPNSSLERYDGGPPFPSCIPRIQNLGLFPFIHGMMWKITPIATRSNRL 1020
Qy 1021 KLRLDLTDEANTSCNIWNGSTBYRPTPVVEEDLRAPQLNHSNSDVVSNCHRLVFNQEI 1080
Db 1021 KLRLDLTDEANTSCNIWNGSTBYRPTPVVEEDLRAPQLNHSNSDVVSNCHRLVFNQEI 1080
Qy 1081 NFHLIGNLWLSRKALKYKMKIMVNAALQROFHSPPIFREDDPSRQI 1128
Db 1081 NFHLIGNLWLSRKALKYKMKIMVNAALQROFHSPPIFREDDPSRQI 1128

RESULT 4
ID AAB50085 standard; protein; 1188 AA.
AC AAB50085;
XX
DT 19-MAR-2001 (first entry)
DE Human A259.
XX
KW Human; A259; integrin alpha subunit; fibrosis; liver disease; lung;
KW kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX

FH Key Location/Qualifiers
FT Domain 1..1141
FT /label= Extracellular_domain
FT Peptide 1..22
FT /label= Signal_peptide
FT Protein 23..1188
FT /label= Mature_protein
FT Domain 39..74
FT /label= Integrin_alphasubunit_repeat_domain_#1
FT Domain 115..157
FT /label= Integrin_alphasubunit_repeat_domain_#2
FT Domain 164..345
FT /label= I_domain
FT Domain 367..392
FT /label= Integrin_alphasubunit_repeat_domain_#3
FT Domain 421..455
FT /label= Integrin_alphasubunit_repeat_domain_#4
FT Domain 478..516
FT /label= Integrin_alphasubunit_repeat_domain_#5
FT Domain 540..575
FT /label= Integrin_alphasubunit_repeat_domain_#6
FT Domain 602..640
FT /label= Integrin_alphasubunit_repeat_domain_#7
FT Domain 1142..1164
FT /label= Transmembrane_domain
FT Domain 1165..1188
FT /label= Cytoplasmic_domain
XX
PN WO200073339-A1.
XX
PD 07-DEC-2000.
XX
PP 15-MAY-2000; 2000WO-US013262.
XX
PR 28-MAY-1999; 99US-00322790.
PR 27-APR-2000; 2000US-00561263.
XX
PA (MILL-) MILLLENNIUM PHARM INC.
XX
PI Pan Y, Lora JW;
XX
DR WPI; 2001-041142/05.
DR N-PSDB; AAC91901, AAC91902.
XX
PT Nucleic acid encoding alpha-integrin subunits, useful for treatment and
diagnosis of fibrosis, e.g. of the liver.

```

```

XX Claim 8; Fig 1; 164pp; English.
PS
CC The present sequence is human integrin alpha subunit. A259. A259 is
CC homologous with the alpha1 and alpha2 integrin subunits and is
CC overexpressed in fibrosis. A259 is implicated in regulation of
CC proliferation, differentiation and/or function of many different cell
CC types. Inhibitors of A259 activity are useful for the treatment of liver
CC disease, particularly fibrosis, and also fibrosis in other organs
CC (specifically lung and kidney). In addition, A259 can be used for
CC treatment and prevention of cancer, osteoporosis, acute myeloid
CC leukaemia, HIV infection, and rheumatoid arthritis
XX
SQ Sequence 1188 AA;
Query Match 77.9%; Score 926; DB 4; Length 1188;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1126; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MDLPRGLVAVWALSINPGFTDTFNMDFKPRVPGSRTAPRGYVQOHDISGNKWLAVGA 60
Db 1 MDLPRGLVAVWALSINPGFTDTFNMDFKPRVPGSRTAPRGYVQOHDISGNKWLAVGA 60
Qy 61 PLENGYQKTGDVYKCPVIHGNCTKLNGLRVTLNSVSRKDNMELGLSLATNPKDNSFLA 120
Db 61 PLENGYQKTGDVYKCPVIHGNCTKLNGLRVTLNSVSRKDNMELGLSLATNPKDNSFLA 120
Qy 121 CSPLWSHECGSSYTTGMCSEKRVNSFRPSKTVAPALQRCQTYMDIVIVLDSNIIYPWVE 180
Db 121 CSPLWSHECGSSYTTGMCSEKRVNSFRPSKTVAPALQRCQTYMDIVIVLDSNIIYPWVE 180
Qy 181 VOHFLINILKKFYIGPGQIQGVVVOYGEDVVHPEHLDYRSVKDVAASHIEORGCTET 240
Db 181 VOHFLINILKKFYIGPGQIQGVVVOYGEDVVHPEHLDYRSVKDVAASHIEORGCTET 240
Qy 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDSEHSDPDLKVKVQOERDNTVYAVAVL 300
Db 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDSEHSDPDLKVKVQOERDNTVYAVAVL 300
Qy 301 GYVNRGINPETFLNEIKYIASDPPDKHFNVTDEAALKDVIDALGDRIFSELTGNKET 360
Db 301 GYVNRGINPETFLNEIKYIASDPPDKHFNVTDEAALKDVIDALGDRIFSELTGNKET 360
Qy 361 SFGLEMSQTFSSHVVEDGVLGAVGAYDMNGAVLKETSAGKVIPLRESYLKSPPELKN 420
Db 361 SFGLEMSQTFSSHVVEDGVLGAVGAYDMNGAVLKETSAGKVIPLRESYLKSPPELKN 420
Qy 421 HGAYLGTVTVSVSSRQGRVYVAGAPRNTGKVIPLTMMNRSLTIHQAMRGQOIGSYF 480
Db 421 HGAYLGTVTVSVSSRQGRVYVAGAPRNTGKVIPLTMMNRSLTIHQAMRGQOIGSYF 480
Qy 481 GSEITSDVDIDGDGVTDLVLLVGAPMYFNEGRERKGVYVYELRQNFVYNGTLKDSHSYQNA 540
Db 481 GSEITSDVDIDGDGVTDLVLLVGAPMYFNEGRERKGVYVYELRQNFVYNGTLKDSHSYQNA 540
Qy 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNAGAIYIFHGFRGSILKTPKORITASELATG 600
Db 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNAGAIYIFHGFRGSILKTPKORITASELATG 600
Qy 601 LOYFGCSIHGOLDLNEGLDLAVGALGNVILMSRPVVOINASLHTEPSPKINIFHRDCK 660
Db 601 LOYFGCSIHGOLDLNEGLDLAVGALGNVILMSRPVVOINASLHTEPSPKINIFHRDCK 660
Qy 661 RSGRDATCLAAFLCFTPIFLAPHQTTTGVIRYNATMDERRYTPRAHLDEGGRFTTRAV 720
Db 661 RSGRDATCLAAFLCFTPIFLAPHQTTTGVIRYNATMDERRYTPRAHLDEGGRFTTRAV 720
Qy 721 LLSSGQELCERINFHVLDTADYKPVTFSEYSELEDDHGPMLDDGPTTLRVSVFPWNG 780
Db 721 LLSSGQELCERINFHVLDTADYKPVTFSEYSELEDDHGPMLDDGPTTLRVSVFPWNG 780
Qy 781 CNEDEHCVDPDLVDARSDLPTAMEYCORVLRKPAQCDSAYTSLSPDTTFFIESTQRVAV 840
Db 781 CNEDEHCVDPDLVDARSDLPTAMEYCORVLRKPAQCDSAYTSLSPDTTFFIESTQRVAV 840

```

781	CNEDEHCVDPDLVDARSDLPATAKEYQORVLRPAQCDSAYTLSFOTTVFIIESTRQRVAV	840
841	EATLENRGENAYSTVLNISQSANLQASLTQKEDSDGSIETCNVEERLQKVCNVSYPFF	900
841	EATLENRGENAYSTVLNISQSANLQASLTQKEDSDGSIETCNVEERLQKVCNVSYPFF	900
901	RAKAKVAFRLDSEFKSILPHHLLEILAAGSDSNERDSTKDENVAPLRFLHKYADVLFT	960
901	RAKAKVAFRLDSEFKSILPHHLLEILAAGSDSNERDSTKDENVAPLRFLHKYADVLFT	960
961	RSSLSHYEVKLNSSLERYDGIQPPSPCIFRIQNGLGFPIHGMMKITIPIATRSGNRLL	1020
961	RSSLSHYEVKPNSSLERYDGIQPPSCIFRQNGLGFPIHGMMKITIPIATRSGNRLL	1020
1021	KLRDFLTDEANTSCNIWGNSTBYRPPTVEDLRRAQLNHNSDDVVSNICNLRLVPNOEI	1080
1021	KLRDFLTDEANTSCNIWGNSTBYRPPTVEDLRRAQLNHNSDDVVSNICNLRLVPNOEI	1080
1081	NFHLLGNLWLRSALKYKSKMIWNAAALQROPHSPFI FREEDPSROI	1128
1081	NFHLLGNLWLRSALKYKSKMIWNAAALQROPHSPFI FREEDPSROI	1128
RESULT 5		
AAU10551		
ID	AAU10551 standard; protein; 1188 AA.	
XX		
AC	AAU10551;	
XX		
DT	14-FEB-2002 (first entry)	
XX		
DE	Fuman A259 polypeptide.	
XX		
KW	Human; A259; integrin alpha subunit; integrin alpha 10; secreted protein; liver disease; fibrosis; lung; kidney; bone associated disorder; blood; cartilage associated disorder; haematopoietic disorder; bone marrow; immune related disease; apoptotic disorder; neuronal tissue disease; neurodegenerative disease; gene therapy; cancer; cytostatic; osteopathic; nephrotropic; immunomodulator; anti-inflammatory; neuroprotective; antiarthritic; antianaemic; anti-allergic; antiasthmatic; dermatological; antidiabetic; anticonvulsant; antiparkinsonian.	
OS	Homo sapiens.	
XX		
Key	Location/Qualifiers	
Domain	1. .1141 /note= "Extracellular domain"	
Peptide	1..22 /note= "Signal peptide"	
Protein	23..1188 /note= "Mature human A259"	
Domain	37. .90 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"	
Domain	115. .157 /note= "Integrin alpha repeat domain"	
Domain	164. .345 /note= "I domain or Von Willebrand Factor type A domain"	
Domain	367. .392 /note= "Integrin alpha repeat domain"	
Domain	421. .472 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"	
Domain	476. .532 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"	
Domain	538. .593 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"	
Domain	600. .654 /note= "Integrin alpha repeat domain. The specification states that this domain exists in human A549"	
Domain	1142. .1164	

QY 361 SFGLMSQTGFSSHVVGDGVLGAYDNGAVLKETTSAGKVIPLRESYLKEPPEELKN 420  
 DB |||||  
 QY 361 SFGLMSQTGFSSHVVGDGVLGAYDNGAVLKETTSAGKVIPLRESYLKEPPEELKN 420  
 DB |||||  
 QY 421 HGAYLGTVTVSVSSRQGRVYVAGAPRPNHETGKVLFTMHNNRSLTIHQAMRGOQIGSYF 480  
 DB |||||  
 QY 421 HGAYLGTVTVSVSSRQGRVYVAGAPRPNHETGKVLFTMHNNRSLTIHQAMRGOQIGSYF 480  
 DB |||||  
 QY 481 GSEITSVDIDGGVTDVLLVGAPMYENEGREGKVVYVELRQNFVYNGTLKDSYQNA 540  
 DB |||||  
 QY 481 GSEITSVDIDGGVTDVLLVGAPMYENEGREGKVVYVELRQNFVYNGTLKDSYQNA 540  
 DB |||||  
 QY 541 RFGSSIASVRDLNODSYNDVVVYGAPLEDNHAGAIYIFHGFSGIILKTPKQRITASELATG 600  
 DB |||||  
 QY 541 RFGSSIASVRDLNODSYNDVVVYGAPLEDNHAGAIYIFHGFSGIILKTPKQRITASELATG 600  
 DB |||||  
 QY 601 LOYFGCSHCOLDAEDGLIDLAVGALGNVILWSPVQINASHFPPSKINIFHRDCK 660  
 DB |||||  
 QY 601 LOYFGCSHCOLDAEDGLIDLAVGALGNVILWSPVQINASHFPPSKINIFHRDCK 660  
 DB |||||  
 QY 661 RSGRATCLAAFLCTPIFLAPHFQTTVIGIRYNATMDERRYTTPRAHLDEGGDRTNRAV 720  
 DB |||||  
 QY 661 RSGRATCLAAFLCTPIFLAPHFQTTVIGIRYNATMDERRYTTPRAHLDEGGDRTNRAV 720  
 DB |||||  
 QY 721 LLSGQELCERINFHVLDTADYVKVPTFSVEYSLEDDPHGPMDDGWPTTLKRVSPFWNG 780  
 DB |||||  
 QY 721 LLSGQELCERINFHVLDTADYVKVPTFSVEYSLEDDPHGPMDDGWPTTLKRVSPFWNG 780  
 DB |||||  
 QY 781 CNEDEHCVPLVDLARSDLPTAMEYCORVLRPAQDCSAYTLSPFTTFIESTRQVAV 840  
 DB |||||  
 QY 781 CNEDEHCVPLVDLARSDLPTAMEYCORVLRPAQDCSAYTLSPFTTFIESTRQVAV 840  
 DB |||||  
 QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSGSIECVNEBRRLQKQVCNVSYPFF 900  
 DB |||||  
 QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSGSIECVNEBRRLQKQVCNVSYPFF 900  
 DB |||||  
 QY 901 RAKAVAPLDEFSKSTPLHLELELAGSDSNERDSTKEDNVAPLPHLYKAEADVLPT 960  
 DB |||||  
 QY 901 RAKAVAPLDEFSKSTPLHLELELAGSDSNERDSTKEDNVAPLPHLYKAEADVLPT 960  
 DB |||||  
 QY 961 RSSLSHVEKLNSSLERYDGIGPPFCIFRIQNLGLPPIHGXMMKITIPIATRSNELL 1020  
 DB |||||  
 QY 961 RSSLSHVEKLNSSLERYDGIGPPFCIFRIQNLGLPPIHGXMMKITIPIATRSNELL 1020  
 DB |||||  
 QY 1021 KLRFELTDEANTSCNMGNSSTERYPTVPEEDLRAPQLNHSNDVVSINCMIRLVPNOEI 1080  
 DB |||||  
 QY 1021 KLRFELTDEANTSCNMGNSSTERYPTVPEEDLRAPQLNHSNDVVSINCMIRLVPNOEI 1080  
 DB |||||  
 QY 1081 NFHLGNLWRLSLKALKYKSMIMVNAALQROFHSPPFIREDPSROI 1128  
 DB |||||  
 QY 1081 NFHLGNLWRLSLKALKYKSMIMVNAALQROFHSPPFIREDPSROI 1128  
 DB |||||  
 RESULT 6  
 AAU14467  
 ID AAU14467 standard; protein; 1188 AA.  
 AC AAU14467;  
 XX AAU14467;  
 XX AAU14467;  
 DT 24-OCT-2001 (first entry)  
 XX Human novel protein #338.  
 DE Human; novel protein; Antianemic; osteopathic; antiinflammatory;  
 KW immunomodulatory; cyrostatic; neuroprotective; vulnerary; nootropic;  
 KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;  
 KW antibacterial; antiallergic; dermatological; haemostatic; antidiabetic;  
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
 KW tissue regeneration; immune disorder.  
 XX Homo sapiens.  
 OS

XX WO200155437-A2.  
 PN  
 XX  
 XX  
 PD 02-AUG-2001.  
 XX  
 PP 25-JAN-2001; 2001WO-US002623.  
 XX  
 PR 25-JAN-2000; 2000US-00491404.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 DR WPI; 2001-451939/48.  
 XX  
 DR N-PSDB; AAS22772.  
 XX  
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,  
 PT nervous system disorders, and for regenerating bone and cartilage.  
 XX  
 PS Example 4; Page 828-831; 894pp; English.  
 XX  
 CC The invention relates to polynucleotides encoding novel human proteins or  
 CC their active domains. The polypeptides, polynucleotides and antibodies  
 CC raised against the polypeptides are used in a method of treatment of a  
 CC mammal and prevention of disorders caused by the aberrant protein  
 CC expression or activity. The polypeptides can be used as molecular weight  
 CC markers, food supplements, and in antibody production. The polypeptides  
 CC are used to identify compounds which bind to the polypeptides.  
 CC Polynucleotides of the invention are used as probes and primers, for  
 CC sequencing, for chromosome or gene mapping, in the production of  
 CC recombinant proteins, and in generating anti-sense DNA or RNA and in gene  
 CC therapy. Polypeptides of the invention can be used to target drugs to a  
 CC tumour, in assays to determine biological activity, to raise  
 CC antibodies/elicite an immune response, to determine quantitative protein  
 CC levels, as tissue markers, and to isolate receptors or ligands.  
 CC Polypeptides of the invention may also be useful in treating platelet  
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
 CC the proliferation, differentiation and survival of stem cells, as a  
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia, as a  
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
 CC fungal infection or from autoimmunity, cancer, allergy, asthma, graft-  
 CC versus-host disease, eczema, haemophilia, thrombosis, anti-inflamatory  
 CC diseases, nervous system disorders, and infection. The present sequence  
 CC represents a protein of the invention  
 XX  
 SQ Sequence 1188 AA;

Query Match 76.7%; Score 911; DB 4; Length 1188;  
 Best Local Similarity 100.0%; Pred.No. 0;  
 Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDLPRGLVVAWALSILWEGFTDTFNMOTRKPVRVPGSRTAFPGYTVQOHDISGNKWLVVGA 60  
 DB 1 MDLPRGLVVAWALSILWEGFTDTFNMOTRKPVRVPGSRTAFPGYTVQOHDISGNKWLVVGA 60  
 QY 61 PLEINGYQKTDVYKCPVHGNCTKMLGRVTLNVSEKDNMELGLSLATNPKDNGSFLA 120  
 DB 61 PLEINGYQKTDVYKCPVHGNCTKMLGRVTLNVSEKDNMELGLSLATNPKDNGSFLA 120  
 QY 121 CSPLWSHECGSSYYTTCMCRVNSNFPFSKTVAPALQRCQTYMDIVLDOGSNIYVWE 180  
 DB 121 CSPLWSHECGSSYYTTCMCRVNSNFPFSKTVAPALQRCQTYMDIVLDOGSNIYVWE 180  
 QY 181 VQHFILNLLKFFYIGRCQIQGVVQYGEDVVHBFHLNDYRSVKDVAASHIEORGGTET 240  
 DB 181 VQHFILNLLKFFYIGRCQIQGVVQYGEDVVHBFHLNDYRSVKDVAASHIEORGGTET 240  
 QY 241 RTAFGIFARSEAFQKGRGAKKVMITVITDGHSDSPDLSEKVIQOSERNVTYAVAVL 300  
 DB 241 RTAFGIFARSEAFQKGRGAKKVMITVITDGHSDSPDLSEKVIQOSERNVTYAVAVL 300

QY 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVNTDEAALKDIDVALGDRIFSLEGTNKNET 360  
DB |||||  
QY 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVNTDEAALKDIDVALGDRIFSLEGTNKNET 360  
DB |||||  
QY 361 SFGLEMSQTGFSSHHVDEGVLLGAVGAYDNWNGAVLKETSAGKVIPLRESYLKEFPBELKN 420  
DB |||||  
QY 361 SFGLEMSQTGFSSHHVDEGVLLGAVGAYDNWNGAVLKETSAGKVIPLRESYLKEFPBELKN 420  
DB |||||  
QY 421 HGAYLGVTYTSVSSRQGRVYVAGAPFRNHTGKVLFTMHNNRSLTIHOAMRGQQIGSYF 480  
DB |||||  
QY 421 HGAYLGVTYTSVSSRQGRVYVAGAPFRNHTGKVLFTMHNNRSLTIHOAMRGQQIGSYF 480  
DB |||||  
QY 481 GSEITSDVDIDGQVTDVLLVAGAPMYFNEGREGKVVYVELRQNRVYVINGTLKDSHSYQNA 540  
DB |||||  
QY 481 GSEITSDVDIDGQVTDVLLVAGAPMYFNEGREGKVVYVELRQNRVYVINGTLKDSHSYQNA 540  
DB |||||  
QY 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFGRGSIILKTPQRITASELATG 600  
DB |||||  
QY 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFGRGSIILKTPQRITASELATG 600  
DB |||||  
QY 601 LOYFGCSIHQDLNEDGLDILAVGALGNVILKSRPVVQINASLHFPESKINIFHRDCK 660  
DB |||||  
QY 601 LOYFGCSIHQDLNEDGLDILAVGALGNVILKSRPVVQINASLHFPESKINIFHRDCK 660  
DB |||||  
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVIGIRYNATMDERRYTTPRAHLDEGGDRFTNRV 720  
DB |||||  
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVIGIRYNATMDERRYTTPRAHLDEGGDRFTNRV 720  
DB |||||  
QY 721 LLSGGQELCRIRNFHVLDTADYVKVPTFSVEYSLEDDPDHGMELDDGWPTTLRVSVFPWNG 780  
DB |||||  
QY 721 LLSGGQELCRIRNFHVLDTADYVKVPTFSVEYSLEDDPDHGMELDDGWPTTLRVSVFPWNG 780  
DB |||||  
QY 781 CNEDEHCVPLVLDARSDLPTAMEYCORVLRKPAQDCSAYTILSPDTTVFIIESTRORVAV 840  
DB |||||  
QY 781 CNEDEHCVPLVLDARSDLPTAMEYCORVLRKPAQDCSAYTILSPDTTVFIIESTRORVAV 840  
DB |||||  
QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERRLQKQVCNVSYPFF 900  
DB |||||  
QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERRLQKQVCNVSYPFF 900  
DB |||||  
QY 901 RAKAKVAFRLD 911  
DB |||||  
QY 901 RAKAKVAFRLD 911  
DB |||||

RESULT 7  
ADE09956  
ID ADE09956 standard; protein; 1188 AA.  
XX AC ADE09956;  
XX DT 29-JAN-2004 (first entry)  
XX DE Novel protein-related contig polypeptide sequence #544.  
XX KW novel gene; novel protein; tissue marker; molecular weight marker;  
XX KM chromosome marker; genetic disorder; contig.  
XX OS Unidentified.  
XX PW W02003054152-A2.  
XX XX 03-JUL-2003.  
XX PD 10-DEC-2002; 2002W0-US039555.  
XX PF 10-DEC-2001; 2001US-0339739P.  
XX PR 11-DEC-2001; 2001US-0339453P.  
XX PR 14-MAR-2002; 2002US-0365091P.  
XX PR 14-MAR-2002; 2002US-0365384P.  
XX PR 12-APR-2002; 2002US-0372381P.  
XX PR 12-APR-2002; 2002US-0372615P.  
XX PR 22-APR-2002; 2002US-00128558.

PR 24-APR-2002; 2002US-0376045P.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Asundi V, Goodrich RN, Ren P, Zhang J, Zhao QA, Wang J;  
XX Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;  
XX Ma Y, Wang D, Chen R, Xu C, Boyle BU;  
XX WPI; 2003-569235/53.  
XX New polynucleotides, useful for expressing recombinant proteins for  
XX analysis, characterization or therapeutic use, or as markers for tissues  
XX in which the corresponding protein is preferentially expressed.  
XX PS Disclosure; SEQ ID NO 3022; 1177bp; English.  
XX CC The invention comprises the amino acid and coding sequences of novel  
XX proteins. The DNA and protein sequences of the invention are useful as:  
XX markers for tissues in which the corresponding protein is preferentially  
XX expressed; as molecular weight markers on gels, as chromosome markers or  
XX tags; to identify chromosomes or to map related gene positions; and to  
XX compare with endogenous DNA sequences in patients to identify potential  
XX genetic disorders. The present amino acid sequence was used in the  
XX exemplification of the invention.  
XX SQ Sequence 1188 AA;  
Query Match 76.7%; Score 911; DB 7; Length 1188;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDLPRLGLVAVALSWPQFTDTFNNMTKPRVPGSRFTAFPGYTVQOHDISGNKWLWVGA 60  
DB 1 MDLPRLGLVAVALSWPQFTDTFNNMTKPRVPGSRFTAFPGYTVQOHDISGNKWLWVGA 60  
QY 61 PLETTYGKTDGVKCPVIHGNCTKLNLRVTLNSVSRKDNMRGLSLATNPKDNSFLA 120  
DB 61 PLETTYGKTDGVKCPVIHGNCTKLNLRVTLNSVSRKDNMRGLSLATNPKDNSFLA 120  
QY 121 CSPLWSHSCGSSYYTTGMSRVNSNFRSKTVAPALQRCQTVMDIVIVLDSNSIYPWVE 180  
DB 121 CSPLWSHSCGSSYYTTGMSRVNSNFRSKTVAPALQRCQTVMDIVIVLDSNSIYPWVE 180  
QY 181 VOHFLINILKRPYIGPGQIQGVGVQYGVGVVHPEHLNDYRSVKDVVEASHIEQGGTET 240  
DB 181 VOHFLINILKRPYIGPGQIQGVGVQYGVGVVHPEHLNDYRSVKDVVEASHIEQGGTET 240  
QY 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITGESHSDSPDLEKVIQOQSRDNVTRVAVVL 300  
DB 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITGESHSDSPDLEKVIQOQSRDNVTRVAVVL 300  
QY 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVNTDEAALKDIDVALGDRIFSLEGTNKNET 360  
DB 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVNTDEAALKDIDVALGDRIFSLEGTNKNET 360  
QY 361 SFGLEMSQTGFSSHHVDEGVLLGAVGAYDNWNGAVLKETSAGKVIPLRESYLKEFPBELKN 420  
DB 361 SFGLEMSQTGFSSHHVDEGVLLGAVGAYDNWNGAVLKETSAGKVIPLRESYLKEFPBELKN 420  
QY 421 HGAYLGVTYTSVSSRQGRVYVAGAPFRNHTGKVLFTMHNNRSLTIHOAMRGQQIGSYF 480  
DB 421 HGAYLGVTYTSVSSRQGRVYVAGAPFRNHTGKVLFTMHNNRSLTIHOAMRGQQIGSYF 480  
QY 481 GSEITSDVDIDGQVTDVLLVAGAPMYFNEGREGKVVYVELRQNRVYVINGTLKDSHSYQNA 540  
DB 481 GSEITSDVDIDGQVTDVLLVAGAPMYFNEGREGKVVYVELRQNRVYVINGTLKDSHSYQNA 540  
QY 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFGRGSIILKTPQRITASELATG 600  
DB 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFGRGSIILKTPQRITASELATG 600  
QY 601 LOYFGCSIHQDLNEDGLDILAVGALGNVILKSRPVVQINASLHFPESKINIFHRDCK 660  
DB |||||

Db 601 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILSRPVVQINASLHFPSPKINIFHRDCK 660  
 QY 661 RSGRDATCLAAFLCTPPIFLAPHFQTTVGIRYNATMBERRYTPRAHLDEGGDRFTNRV 720  
 Db 661 RSGRDATCLAAFLCTPPIFLAPHFQTTVGIRYNATMBERRYTPRAHLDEGGDRFTNRV 720  
 QY 721 LLSGQELCERINFHLVDADYVVKVTSVEYSLEDDPHGMLDGPMTTLRVSVPPFWNG 780  
 Db 721 LLSGQELCERINFHLVDADYVVKVTSVEYSLEDDPHGMLDGPMTTLRVSVPPFWNG 780  
 QY 781 CNEDEHCVPLDVLARSDLPTAMEYCORVLRKPAODCSAYTLSPTTVPILLESFQRVAV 840  
 Db 781 CNEDEHCVPLDVLARSDLPTAMEYCORVLRKPAODCSAYTLSPTTVPILLESFQRVAV 840  
 QY 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGIECVNEERRLQKQVCNVSYPFF 900  
 Db 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGIECVNEERRLQKQVCNVSYPFF 900  
 QY 901 RAKAKVAPRLD 911  
 Db 901 RAKAKVAPRLD 911

## RESULT 8

AAB25582  
 ID AAB25582 standard; protein; 1189 AA.

XX AAB25582;

DT 21-NOV-2000 (first entry)

XX ITGA11 protein encoded by human secreted protein gene #7.

DE Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;  
 KW antirheumatic, dermatological; antiproliferative; antiarteriosclerotic;  
 KW anticancer; vulnary; antiviral; antibacterial; antifungal;  
 KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;  
 KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;  
 KW Crohn's disease; nephritis; hyperproliferative disorder;  
 KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;  
 KW melanoma; lymphoma; wound healing; human; chromosome 15 q22.3-23.

XX Homo sapiens.

XX WO200029435-A1.

XX 25-MAY-2000.

XX 27-OCT-1999; 99WO-US025031.

XX 28-OCT-1998; 98US-0105971P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;

PI Greene JM;

DR WPI; 2000-387742/33.

XX N-PSDB; AAB80612.

XX Isolated nucleic acid molecules encoding human secreted proteins are used  
 PT for the prevention, amelioration and treatment of autoimmune,  
 PT inflammatory, hyperproliferative and cardiovascular disorders, cancer,  
 PT wounds, and infectious diseases.

XX Claim 1; Fig 19A-F; 803pp; English.

XX The present invention relates to 12 secreted human proteins and the  
 CC nucleotide sequences encoding them. The polynucleotide sequences given in  
 CC AAB80606-A80623 encode the 12 secreted protein sequences given in  
 CC AAB25576-B25593. The human secreted proteins have various activities  
 CC dependent on the tissues in which they are expressed. Examples of the  
 CC activities of the proteins include: immunosuppressant; anti-inflammatory;

CC antiarthritic; antirheumatic, dermatological; antiproliferative;  
 CC antiarteriosclerotic; anticancer; vulnary; antiviral; antibacterial;  
 CC and antifungal activity. The proteins, polypeptides, agonists and  
 CC antagonists may be used to treat prevent and/or diagnose various disease,  
 CC disorders and conditions examples of which include: immune disorders e.g.  
 CC Addison's disease, rheumatoid arthritis, dermatitis, and multiple  
 CC sclerosis; inflammatory disorders e.g. inflammatory bowel disease,  
 CC Crohn's disease and nephritis; hyperproliferative disorders such as  
 CC paraproteinemia and purpura; cardiovascular disorders e.g. coronary  
 CC arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The  
 CC proteins and polynucleotide sequences may also be used in wound healing  
 CC and the treatment of infectious diseases. The human secreted protein gene  
 CC #7 and protein sequences are represented in sequences AAB80612 and  
 CC AAB25582. Secreted protein gene #7 is located at position chromosome 15  
 CC q22.3-23. Sequences AAB80652-A80661 represent genes which are related to  
 CC the secreted protein gene#7

XX Sequence 1189 AA;

Query Match 76.7%; Score 911; DB 3; Length 1189;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLPRGLVAVWALSLLWPGFTDTFNMTRKPRVTPGSRAPFGYTVQOHDISGNKWLVVGA 60  
 Db 1 MDLPRGLVAVWALSLLWPGFTDTFNMTRKPRVTPGSRAPFGYTVQOHDISGNKWLVVGA 60  
 QY 61 PLETNQYQKTDVYKCPVHGNCTKLNLRVTLSNVSEKXDMELGSLATNPKDNSPLA 120  
 Db 61 PLETNQYQKTDVYKCPVHGNCTKLNLRVTLSNVSEKXDMELGSLATNPKDNSPLA 120  
 QY 121 CSPLWSHECCSSYYTTCMSRVNSNFRPSKTVAPALQRCOTYMDIVILVDCGNSIYPWE 180  
 Db 121 CSPLWSHECCSSYYTTCMSRVNSNFRPSKTVAPALQRCOTYMDIVILVDCGNSIYPWE 180  
 QY 181 VQFLINILKKFYIGPQIQVGVVQYGEDVVEHFLANDYRSVKDWEAAASHIEQGGTET 240  
 Db 181 VQFLINILKKFYIGPQIQVGVVQYGEDVVEHFLANDYRSVKDWEAAASHIEQGGTET 240  
 QY 241 RTAFGIEFARSEAPQKGRGAKKWMIVITDGESHSDPDLKVIQOESRDNVTYAVAVL 300  
 Db 241 RTAFGIEFARSEAPQKGRGAKKWMIVITDGESHSDPDLKVIQOESRDNVTYAVAVL 300  
 QY 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVTDEAALKDIDVALGDRIIFSLGNTNKNET 360  
 Db 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVTDEAALKDIDVALGDRIIFSLGNTNKNET 360  
 QY 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLMKEPPELKN 420  
 Db 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLMKEPPELKN 420  
 QY 421 HGAYLGVTVTSVSSROGRVYVAGAPRFNHTGKVIPLFTMNNRSLTTHQAWRGQOIGSYP 480  
 Db 421 HGAYLGVTVTSVSSROGRVYVAGAPRFNHTGKVIPLFTMNNRSLTTHQAWRGQOIGSYP 480  
 QY 481 GSEITTSVDIDGDGVTDVLLGAPMYFNEGRERGVYVYELRQNFVYNGTLKDSHSYQNA 540  
 Db 481 GSEITTSVDIDGDGVTDVLLGAPMYFNEGRERGVYVYELRQNFVYNGTLKDSHSYQNA 540  
 QY 541 RFGSSIASVSDLNQDSYNDVVGAPLEDNAGAIYIHPGGRGSLTKPKORTASELATG 600  
 Db 541 RFGSSIASVSDLNQDSYNDVVGAPLEDNAGAIYIHPGGRGSLTKPKORTASELATG 600  
 QY 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILSRPVVQINASLHFPSPKINIFHRDCK 660  
 Db 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILSRPVVQINASLHFPSPKINIFHRDCK 660  
 QY 661 RSGRDATCLAAFLCTPPIFLAPHFQTTVGIRYNATMBERRYTPRAHLDEGGDRFTNRV 720  
 Db 661 RSGRDATCLAAFLCTPPIFLAPHFQTTVGIRYNATMBERRYTPRAHLDEGGDRFTNRV 720  
 QY 721 LLSGQELCERINFHLVDADYVVKVTSVEYSLEDDPHGMLDGPMTTLRVSVPPFWNG 780

Db 721 LLSGQELCERINHFVLDADYKVPVTSVEYSLEDDPDDGPMDDGWPTTLRVSVFPWNG 780

Qy 781 CNEDEHCVPDLVLDASDLPTAMEYCORVLRKPAQCSAYTISFDITVPIIESTRQAV 840

Db 781 CNEDEHCVPDLVLDASDLPTAMEYCORVLRKPAQCSAYTISFDITVPIIESTRQAV 840

Qy 841 EATLENGENAYSTVLNISOSANLQFASLIQKEDSDGSEECVNEERLQKQCNVSPYFF 900

Db 841 EATLENGENAYSTVLNISOSANLQFASLIQKEDSDGSEECVNEERLQKQCNVSPYFF 900

Qy 901 RAKAKVAFLRD 911

Db 901 RAKAKVAFLRD 911

RESULT 9

ABRS8364

ID ABR58364 standard; protein; 1189 AA.

XX AC ABR58364;

XX DT 07-JUL-2003 (first entry)

XX DE Human NOV2a.

XX KW Human; NOV; antidiabetic; anorectic; antibacterial; virucide; immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia; antiparkinsonian; antilipaeamic; gene therapy; metabolic disorder; diabetes; obesity; infection; cachexia; cancer; Parkinson's disease; neurodegenerative disorder; Alzheimer's disease; immune disorder; haematopoietic disorder.

XX OS Homo sapiens.

XX PN WO2003029423-A2.

XX PD 10-APR-2003.

XX PF 02-OCT-2002; 2002WO-US031358.

XX PR 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327342P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 22-OCT-2001; 2001US-0341058P.

PR 24-OCT-2001; 2001US-0339266P.

PR 24-OCT-2001; 2001US-0343829P.

PR 29-OCT-2001; 2001US-0349575P.

PR 01-NOV-2001; 2001US-0346357P.

PR 12-APR-2002; 2002US-0371972P.

PR 17-APR-2002; 2002US-0371980P.

PR 19-APR-2002; 2002US-0373261P.

PR 23-APR-2002; 2002US-0373805P.

PR 16-MAY-2002; 2002US-0381101P.

PR 17-MAY-2002; 2002US-0381635P.

PR 29-MAY-2002; 2002US-0383830P.

PR 01-OCT-2002; 2002US-00262839.

XX (CURA-) CURAGEN CORP.

XX PA Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E; Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W; Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK; Rothenberg ME, Shimkets RA, Smithson G, Spyttek KA, Taupier RJ; Vernet CAM, Voss EZ, Zerbussen BD, Zhong M;

XX WPI; 2003-381625/36.

DR N-PSDB; ACC72076.

XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or dyslipidaemia, and in chromosome mapping, tissue typing or pharmacogenomics.

XX PT

XX PS Claim 1; Page 105; 487pp; English.

XX CC The present invention relates to novel human NOV proteins and their coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV proteins are useful in manufacturing a method for treating a syndrome associated with a human disease. The NOV proteins and coding sequences may be used to diagnose, treat or prevent metabolic disorders such as diabetes or obesity, infections, cachexia, cancer, neurodegenerative disorders such as Alzheimer's disease or Parkinson's disease, immune disorders, haematopoietic disorders and various dyslipidaemias

XX SQ Sequence 1189 AA;

Query Match 76.7%; Score 911; DB 6; Length 1189;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLPRGLVVAWALSILWPGFTDTFNMDTKPRVPGSRTAFPGYTVQOHDISGNKWLVGGA 60

Db 1 MDLPRGLVVAWALSILWPGFTDTFNMDTKPRVPGSRTAFPGYTVQOHDISGNKWLVGGA 60

QY 61 PLETVNGYQKTDGVYKCPVIHGNCTKLNLRVTLSNVSEKKNMRLGLSLATNPKDNSFLA 120

Db 61 PLETVNGYQKTDGVYKCPVIHGNCTKLNLRVTLSNVSEKKNMRLGLSLATNPKDNSFLA 120

QY 121 CSPLMSHECGSSYYTTGMCNRVNSNFRFSKTVAPALQRCQTYMDIVIVLDSGNSIYPWVE 180

Db 121 CSPLMSHECGSSYYTTGMCNRVNSNFRFSKTVAPALQRCQTYMDIVIVLDSGNSIYPWVE 180

QY 181 VOHEFLINILKKFYICPGQIQGVVQYGEDVHVEHFLNDYRSVKDVAASHIEQGGTET 240

Db 181 VOHEFLINILKKFYICPGQIQGVVQYGEDVHVEHFLNDYRSVKDVAASHIEQGGTET 240

QY 241 RTAFGIEFARSEAFQKGRKGAKKVMIITDGSCHSDSPLEKVIQSSERDNTVRYAVAVL 300

Db 241 RTAFGIEFARSEAFQKGRKGAKKVMIITDGSCHSDSPLEKVIQSSERDNTVRYAVAVL 300

QY 301 GYNNRRGINPETFLNEIKYIASDDDDKHFNNVTDEAALKDIDVALGDRIFSLGCTNKNET 360

Db 301 GYNNRRGINPETFLNEIKYIASDDDDKHFNNVTDEAALKDIDVALGDRIFSLGCTNKNET 360

QY 361 SFGLEMSQTFSSHVVEDGVLGAVGYDMNGAVLKETSAKVIPLRESYLKEPPEELKN 420

Db 361 SFGLEMSQTFSSHVVEDGVLGAVGYDMNGAVLKETSAKVIPLRESYLKEPPEELKN 420

QY 421 HGAYLGTYTTSVSSRQGRVYVAGAPRNHTGKVLFTMHNNRSLTIHQANRGOQIGSYF 480

Db 421 HGAYLGTYTTSVSSRQGRVYVAGAPRNHTGKVLFTMHNNRSLTIHQANRGOQIGSYF 480

QY 481 GSEITTSVDIDGVTDLVLLVGAPVYFNEGRGRKGVVYVELQRNRPVYNGTLKDSHSYQNA 540

Db 481 GSEITTSVDIDGVTDLVLLVGAPVYFNEGRGRKGVVYVELQRNRPVYNGTLKDSHSYQNA 540

QY 541 RFGSSIASVRDLNQSDYNDVVGAPLEDNHAGAIYIHHGFRGSLKTPKQKITASELATG 600

Db 541 RFGSSIASVRDLNQSDYNDVVGAPLEDNHAGAIYIHHGFRGSLKTPKQKITASELATG 600

QY 601 LOYFGCSITHGOLDLNEEDGLIDIAVGALGNVILSRPVQINASLHPPSPKINIFHRDCK 660

Db 601 LOYFGCSITHGOLDLNEEDGLIDIAVGALGNVILSRPVQINASLHPPSPKINIFHRDCK 660

QY 661 RSGRDATCLAAFLCFTPIFLAPHFTQTTVIGIRYNATMDERRYTPRAHLDEGGDRTNRAV 720

Db 661 RSGRDATCLAAFLCFTPIFLAPHFTQTTVIGIRYNATMDERRYTPRAHLDEGGDRTNRAV 720

QY 721 LLSGQELCERINHFVLDADYKVPVTSVEYSLEDDPDDGPMDDGWPTTLRVSVFPWNG 780

Db 721 LISSGOELCERINFHVLADYVYKPTFSVYSLEDPDRGMLDGGWPTLSVSPFWNG 780  
QY 781 CNEDEHCVPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTSLSDTTTFVFIESTRQVAV 840  
Db 781 CNEDEHCVPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTSLSDTTTFVFIESTRQVAV 840  
QY 841 EATLNRGNAYSTVLNISQSANLPASLIQKEDSDGSGTECNERRRLOKQVCNVSYPFF 900  
Db 841 EATLNRGNAYSTVLNISQSANLPASLIQKEDSDGSGTECNERRRLOKQVCNVSYPFF 900  
QY 901 RAKAKVAFRLD 911  
Db 901 RAKAKVAFRLD 911  
RESULT 10  
ADA27054  
ID ADA27054 standard; protein; 1189 AA.  
XX ADA27054;  
AC ADA27054;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Human novel secreted protein from cDNA HOHBY69 #1.  
XX  
KW cytostatic; antiinflammatory; immunomodulator; neuroprotective;  
KW hemostatic; gene therapy; cancer; inflammation; immune disorder;  
KW neurological disorder; blood clotting disorder; food additive;  
KW preservative; human; secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX US2003055231-A1.  
EN  
XX  
PD 20-MAR-2003.  
XX  
PF 29-OCT-2001; 2001US-00984130.  
XX  
PR 28-OCT-1998; 98US-0105971P.  
PR 27-OCT-1999; 99KO-US025031.  
PR 19-APR-2000; 2000US-0198407P.  
PR 30-OCT-2000; 2000US-0243792P.  
PR 18-APR-2001; 2001US-00836353.  
XX  
PA (NIJ/) NI J.  
PA (YOUN/) YOUNG P B.  
PA (KERN/) KENNY J J.  
PA (OLSE/) OLSEN H S.  
PA (MOOR/) MOORE P A.  
PA (WEIY/) WEI Y.  
PA (GREG/) GREENE J W.  
PA (RUBE/) RUBEN S M.  
PA (LIUD/) LIU D.  
PA (CROC/) CROCKER P R.  
XX  
PI Ni J, Young PB, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;  
PI Ruben SM, Liu D, Crocker PR;  
XX  
DR WPI; 2003-567103/53.  
DR N-PSDB; ADA27036.  
XX  
PT New human secreted nucleic acid molecules and polypeptides, useful for  
PT preventing, treating, or ameliorating a medical condition, such as  
PT cancer, inflammation, immune disorders, neurological and blood clotting  
PT disorders.  
XX  
PS Claim 11; Fig 19; 454pp; English.  
XX  
XX The invention relates to an isolated nucleic molecule that is at least  
CC 95% identical to 18 human cDNA sequences representing 12 novel genes  
CC encoding secreted proteins or a polynucleotide fragment of the cDNA  
CC sequence contained in American Type Culture Collection (ATCC) deposit No.

CC defined in the specification, its species homologue, a variant or allelic  
CC variant of the polynucleotide having a polynucleotide capable of  
CC hybridising under conditions the polynucleotide, where the polynucleotide  
CC does not hybridise under stringent conditions to a nucleic acid molecule  
CC having a nucleotide sequence of only A or T residues. Also included are  
CC recombinant vectors, host cells (for producing the polypeptide), the  
CC secreted polypeptide (comprising a sequence that is at least 95%  
CC identical to a polypeptide fragment, domain, epitope, full-length  
CC protein, variant, allelic variant or species homologue), antibodies that  
CC specifically bind to the polypeptides, diagnosing, treating, preventing  
CC or ameliorating a medical condition by administering the polynucleotide  
CC or the polypeptide, the gene corresponding to the cDNA sequence and  
CC identifying an activity in a biological assay (by expressing the cDNA  
CC sequence in a cell, isolating the supernatant, and detecting an activity  
CC in a biological assay and identifying the protein in the supernatant  
CC having the activity). The polypeptides, nucleic acids and antibodies are  
CC useful for diagnosing a pathological condition or a susceptibility to a  
CC pathological condition, for preventing, treating, or ameliorating a  
CC medical condition, such as cancer, inflammation and other immune  
CC disorders, neurological and blood clotting disorders (many examples are  
CC given in the specification). The nucleic acids are also useful for  
CC chromosome identification, radiation hybrid mapping or long-range  
CC restriction mapping. The polypeptides and antibodies are useful for  
CC providing immunological probes for differential identification of the  
CC tissues immunohistochemistry assays. The polypeptide, polynucleotide,  
CC agonist or antagonist may also be used as a food additive or preservative  
CC to increase or decrease storage capabilities, fat content or other  
CC nutritional components. The present is a secreted protein of the  
CC invention.  
XX  
SQ Sequence 1189 AA;  
Query Match 76.7%; Score 911; DB 6; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MDLPRGLVVAWALSILWFGFTDTPNMDTRKPRVPGSRTAPFGYTVQQHDSIGNKWLVVGA 60  
Db 1 MDLPRGLVVAWALSILWFGFTDTPNMDTRKPRVPGSRTAPFGYTVQQHDSIGNKWLVVGA 60  
QY 61 PLETNGYQKTDGVYKCPVHGNCTKLNIGRTLSNVSEKDNMRGLSLATNPKNDSFLA 120  
Db 61 PLETNGYQKTDGVYKCPVHGNCTKLNIGRTLSNVSEKDNMRGLSLATNPKNDSFLA 120  
QY 121 CSPLWSHECGSSYTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDCGNSIYPWVE 180  
Db 121 CSPLWSHECGSSYTTGMCSSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDCGNSIYPWVE 180  
QY 181 VQHFLINILKFFYTGPGQIQGVVQYGEDVVEHFLNDYRSKDVVEAASHIEORGTTET 240  
Db 181 VQHFLINILKFFYTGPGQIQGVVQYGEDVVEHFLNDYRSKDVVEAASHIEORGTTET 240  
QY 241 RTAFGEIFARSEAFQKGRGKAKKVMIVITDGRSHSDPDLKVIQQSERDNVTRYAVAVL 300  
Db 241 RTAFGEIFARSEAFQKGRGKAKKVMIVITDGRSHSDPDLKVIQQSERDNVTRYAVAVL 300  
QY 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKOIVDALGDRIFSLGNTNNKET 360  
Db 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKOIVDALGDRIFSLGNTNNKET 360  
QY 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDNGAVLKETSGAKVPLRESYLKEPPEELKN 420  
Db 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDNGAVLKETSGAKVPLRESYLKEPPEELKN 420  
QY 421 HGAYLGYTVTWSVSSRQGRVYVAGAPRNFHTGKVILFTMNNRSLTTHQAMRGOQIGSYF 480  
Db 421 HGAYLGYTVTWSVSSRQGRVYVAGAPRNFHTGKVILFTMNNRSLTTHQAMRGOQIGSYF 480  
QY 481 GSEITTSVDIDGVTVDLLVGPAMYFNEGRERGVVYVVELRQNFVYNGTLKDSHSYQNA 540  
Db 481 GSEITTSVDIDGVTVDLLVGPAMYFNEGRERGVVYVVELRQNFVYNGTLKDSHSYQNA 540  
QY 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHDAGAIYIFHGFRGSILKTPKORITASELATG 600



Db 541 RFGSSIASVRLNODSNDVVVCGAPLEDNHAGAIYIFHGRGSIKATPKQRTASELATG 600  
Qy 601 LOYFGCSIHGQDLNEDGLDLAVGALGNVILWSRPVQVQINASLHFPESKINIFHRDCK 660  
Db 601 LOYFGCSIHGQDLNEDGLDLAVGALGNVILWSRPVQVQINASLHFPESKINIFHRDCK 660  
Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720  
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720  
Qy 721 LISSGOELCBRIHFVLDATDYKPTVFSVEYSLDDPDRGPMDDGWPTTLAVSYPFWNG 780  
Db 721 LISSGOELCBRIHFVLDATDYKPTVFSVEYSLDDPDRGPMDDGWPTTLAVSYPFWNG 780  
Qy 781 CNEDEHCVPLVLDARSDLPTAMEYQORVLRKPAQCSAVTSLFSDTFTVFIIESTRORVAV 840  
Db 781 CNEDEHCVPLVLDARSDLPTAMEYQORVLRKPAQCSAVTSLFSDTFTVFIIESTRORVAV 840  
Qy 841 EATLERNGENAYSTVLNISOSANLQFASLLOKEDSDGSIETCNEERLQKQVNSYPPF 900  
Db 841 EATLERNGENAYSTVLNISOSANLQFASLLOKEDSDGSIETCNEERLQKQVNSYPPF 900  
Qy 901 RAKAKVAFLRD 911  
Db 901 RAKAKVAFLRD 911

RESULT 11  
ADE63570  
ID ADE63570 standard; protein; 1189 AA.  
XX AC ADE63570;  
XX DT 29-JAN-2004 (first entry)  
XX DE Human Protein Q9URX5, SEQ ID NO 9514.  
XX KW Human; pain; neuronal tissue; gene therapy;  
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;  
XX KW spared nerve injury; SNI; Chung.  
XX OS Homo sapiens.  
XX PN WO2003015475-A2.  
XX PD 27-FEB-2003.  
XX PF 14-AUG-2002; 2002WO-US025765.  
XX PR 14-AUG-2001; 2001US-0312147P.  
XX PR 01-NOV-2001; 2001US-0346382P.  
XX PR 26-NOV-2001; 2001US-0333347P.  
XX PA (GEHO ) GEN HOSPITAL CORP.  
XX PA (FARB ) BAYER AG.  
XX PI Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
XX PT New composition comprising two or more isolated polypeptides, useful for  
XX PT preparing a medicament for treating pain in an animal.  
XX PS Claim 1; Page: 1017pp; English.  
XX CC The invention discloses a composition comprising two or more isolated rat  
XX CC or human polynucleotides or a polynucleotide which represents a fragment,  
XX CC derivative or allelic variation of the nucleic acid sequence. Also  
XX CC claimed are a vector comprising the novel polynucleotide, a host cell  
XX CC comprising the vector, a method for identifying a nucleotide sequence  
XX CC which is differentially regulated in an animal subjected to pain and a  
XX CC kit to perform the method, an array, a method for identifying an agent

CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 1189 AA;  
Query Match 76.7%; Score 911; DB 7; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDLPRLGVAVALSILWPGFTDTFNMDFKPRVIFGSRTPAFGTYVQOHDISGNKVLVGA 60  
Db 1 MDLPRLGVAVALSILWPGFTDTFNMDFKPRVIFGSRTPAFGTYVQOHDISGNKVLVGA 60  
Qy 61 PLETVGYQKTDVYKCPVIHGNCTKLNLRGVTLSNVSRKDNMRGLSLATNPXNSFLA 120  
Db 61 PLETVGYQKTDVYKCPVIHGNCTKLNLRGVTLSNVSRKDNMRGLSLATNPXNSFLA 120  
Qy 121 CSPLMSHECGSSYTTGMCNRVNFPSKTVAPALQRCQTYMDIVILVDSNSNYPWVE 180  
Db 121 CSPLMSHECGSSYTTGMCNRVNFPSKTVAPALQRCQTYMDIVILVDSNSNYPWVE 180  
Qy 181 VQHFLINILKIFYIGPGQIQGVVQYGEDVHVEFLNDYRSVKDVEAASHIEQGGTET 240  
Db 181 VQHFLINILKIFYIGPGQIQGVVQYGEDVHVEFLNDYRSVKDVEAASHIEQGGTET 240  
Qy 241 RTAFGIEPARSEAFQKGRKCAKVMIVITDGEHSDSPLEKVIQSSERDNNVRYAVAVL 300  
Db 241 RTAFGIEPARSEAFQKGRKCAKVMIVITDGEHSDSPLEKVIQSSERDNNVRYAVAVL 300  
Qy 301 GYNNRRGINPETELNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSLGNTKNET 360  
Db 301 GYNNRRGINPETELNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSLGNTKNET 360  
Qy 361 SFGLEMSQTGFSSHVEDGVLLGAVGAYDMNGAVLKETSAGKVIPLRESYLKBPPEELKN 420  
Db 361 SFGLEMSQTGFSSHVEDGVLLGAVGAYDMNGAVLKETSAGKVIPLRESYLKBPPEELKN 420  
Qy 421 HGAVLGYTVTSVSSRQGRVYVAGAPRNHVGKVLFTMHNNRSITTHOMRGQOIGSYF 480  
Db 421 HGAVLGYTVTSVSSRQGRVYVAGAPRNHVGKVLFTMHNNRSITTHOMRGQOIGSYF 480  
Qy 481 GSEITSDVIDGEGYTDVLLVGAPMYFNEGRERGVYVYELFRQNFVYNGTLKDSHYQVA 540  
Db 481 GSEITSDVIDGEGYTDVLLVGAPMYFNEGRERGVYVYELFRQNFVYNGTLKDSHYQVA 540  
Qy 541 RFGSSIASVRLNODSNDVVVCGAPLEDNHAGAIYIFHGRGSIKATPKQRTASELATG 600  
Db 541 RFGSSIASVRLNODSNDVVVCGAPLEDNHAGAIYIFHGRGSIKATPKQRTASELATG 600  
Qy 601 LOYFGCSIHGQDLNEDGLDLAVGALGNVILWSRPVQVQINASLHFPESKINIFHRDCK 660  
Db 601 LOYFGCSIHGQDLNEDGLDLAVGALGNVILWSRPVQVQINASLHFPESKINIFHRDCK 660  
Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720  
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720



Db 661 RSGRDATCLAAFLCTPIFLAPHFOTTTVIGIRYNATMDERRYTTPRAHLDEGGDRFTNRV 720  
Qy 721 LLSGQELCERINFLVLDADVVKPVTSEVSEYSLSDPDHGMDDGMFTTLRVSVFVWG 780  
Db 721 LLSGQELCERINFLVLDADVVKPVTSEVSEYSLSDPDHGMDDGMFTTLRVSVFVWG 780  
Qy 781 CNEDEHCVDFDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTTLSDFTTTFIESTRQVAV 840  
Db 781 CNEDEHCVDFDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTTLSDFTTTFIESTRQVAV 840  
Qy 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGSEICVNEERLQKQCNVSYPPF 900  
Db 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGSEICVNEERLQKQCNVSYPPF 900  
Qy 901 RAKAKVAFRLD 911  
Db 901 RAKAKVAFRLD 911

RESULT 12  
ADE86584  
ID ADE86584 standard; protein; 1189 AA.

XX ADE86584;  
AC ADE86584;  
XX 29-JAN-2004 (first entry)  
DT Novel human secreted protein #7.  
DE human; secreted protein; cancer; liver disorder; hepatitis;  
KW neural disorder; Alzheimer's disease.  
XX Homo sapiens.

XX US2003129685-A1.  
PN 10-JUL-2003.  
XX 18-APR-2001; 2001US-00836353.  
XX 28-OCT-1998; 98US-0105971P.  
PR 27-OCT-1999; 99WO-US025031.  
PR 19-APR-2000; 2000US-0198407P.

XX (N1J/) NI J.  
PA (YOUN/) YOUNG P E.  
PA (KERN/) KENNY J J.  
PA (OLSE/) OLSEN H S.  
PA (MOOR/) MOORE P A.  
PA (WEI/) WEI Y.  
PA (GRE/) GREENE J M.  
PA (RUB/) RUBEN S M.

XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;  
PI Ruben SM;  
XX WPI: 2004-020335/02.  
DR N-PSDB; AD86566.

XX New nucleic acid molecule, useful for preparing a medicament for  
PT preventing, treating or ameliorating a medical condition e.g. cancer,  
PT liver disorders or neural disorders.

XX Claim 11; SEQ ID NO 35; 380pp; English.

XX The invention relates to an isolated nucleic acid sequence, or its  
CC allelic variant, a fragment of the cDNA sequence, or its fragment,  
CC domain, epitope or species homologue. The nucleic acid is useful for  
CC preparing a medicament for preventing, treating or ameliorating a medical  
CC condition e.g., cancer, liver disorders such as hepatitis or neural  
CC disorders such as Alzheimer's disease. The present sequence represents  
CC the amino acid sequence of a novel human secreted protein.

SQ Sequence 1189 AA;  
Query Match 76.7%; Score 911; DB 8; Length 1189;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDLPRGLVAVWALSFWPGFTDTFNMDTRKPRVIPSRTAFPGYTVOQHDIQGNKWLWVGA 60  
Db 1 MDLPRGLVAVWALSFWPGFTDTFNMDTRKPRVIPSRTAFPGYTVOQHDIQGNKWLWVGA 60  
Qy 61 PLENGYQKTDVYKCPVIHGNCTKLNLRVTLSNVERKDNMRGLSLATNPKNDSFLA 120  
Db 61 PLENGYQKTDVYKCPVIHGNCTKLNLRVTLSNVERKDNMRGLSLATNPKNDSFLA 120  
Qy 121 CSPLWSEHCSSYYTTGMCNRVNSNFRFSKTVAPALQRCQTYMDIVIVLDCSNSIYPWVE 180  
Db 121 CSPLWSEHCSSYYTTGMCNRVNSNFRFSKTVAPALQRCQTYMDIVIVLDCSNSIYPWVE 180  
Qy 181 VQHFLINILKFPYIGPGQIQGVVQYGEDVWVHEFLNDYRSVKDVAASHIEORGCTET 240  
Db 181 VQHFLINILKFPYIGPGQIQGVVQYGEDVWVHEFLNDYRSVKDVAASHIEORGCTET 240  
Qy 241 RTAFGIIFARSEAFQKGRKGAKKVMIVITDGEHSDSPDLEKVIQOESRDNVTVYAVAVL 300  
Db 241 RTAFGIIFARSEAFQKGRKGAKKVMIVITDGEHSDSPDLEKVIQOESRDNVTVYAVAVL 300  
Qy 301 GYNRRGINPETFLNEIKYIASDDPKHFNVTDEAALKOIVDALGDRIFSLGTNKHET 360  
Db 301 GYNRRGINPETFLNEIKYIASDDPKHFNVTDEAALKOIVDALGDRIFSLGTNKHET 360  
Qy 361 SFGLEMSQTGFSSHVEDGVJLGAVGAYDNAGVLTAKTSAGKVIPLRESYLKEPPEELKN 420  
Db 361 SFGLEMSQTGFSSHVEDGVJLGAVGAYDNAGVLTAKTSAGKVIPLRESYLKEPPEELKN 420  
Qy 421 HGAYLGYTTSVSSRQGRVYVAGAPRPHNTGKVLFTMHNNRSLTIHQAMRGOIGSYF 480  
Db 421 HGAYLGYTTSVSSRQGRVYVAGAPRPHNTGKVLFTMHNNRSLTIHQAMRGOIGSYF 480  
Qy 481 GSEITSVDIDGQVTDVLLVCAPMYFNEGRERGVYVYELQNRFRVYNGTLKDSHSYQNA 540  
Db 481 GSEITSVDIDGQVTDVLLVCAPMYFNEGRERGVYVYELQNRFRVYNGTLKDSHSYQNA 540  
Qy 541 RFGSSIASVRDLNDSYNDVVVVGAPLEDNHAGAIYIIFHGFRGSILKTPKQRTITASELATG 600  
Db 541 RFGSSIASVRDLNDSYNDVVVVGAPLEDNHAGAIYIIFHGFRGSILKTPKQRTITASELATG 600  
Qy 601 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILSRPVQINASLHPEPSKINIFHRDCK 660  
Db 601 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILSRPVQINASLHPEPSKINIFHRDCK 660  
Qy 661 RSGRDATCLAAFLCTPIFLAPHFOTTTVIGIRYNATMDERRYTTPRAHLDEGGDRFTNRV 720  
Db 661 RSGRDATCLAAFLCTPIFLAPHFOTTTVIGIRYNATMDERRYTTPRAHLDEGGDRFTNRV 720  
Qy 721 LLSGQELCERINFLVLDADVVKPVTSEVSEYSLSDPDHGMDDGMFTTLRVSVFVWG 780  
Db 721 LLSGQELCERINFLVLDADVVKPVTSEVSEYSLSDPDHGMDDGMFTTLRVSVFVWG 780  
Qy 781 CNEDEHCVDFDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTTLSDFTTTFIESTRQVAV 840  
Db 781 CNEDEHCVDFDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTTLSDFTTTFIESTRQVAV 840  
Qy 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGSEICVNEERLQKQCNVSYPPF 900  
Db 841 EATLNRGENAYSTVLNISQSANLQFASLIQKEDSDGSEICVNEERLQKQCNVSYPPF 900  
Qy 901 RAKAKVAFRLD 911  
Db 901 RAKAKVAFRLD 911

RESULT 13  
RAB25590

ID AAB25590 standard; protein; 1034 AA.

AC AAB25590;

DT 21-NOV-2000 (first entry)

XX Protein encoded by human secreted protein gene #7 clone HOHB69.

XX Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;

KW antineumatic; dermatological; antiproliferative; antiarteriosclerotic;

KW anticancer; vulnery; antiviral; antibacterial; antifungal;

KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;

KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;

KW Crohn's disease; nephritis; hyperproliferative disorder;

KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;

KW melanoma; lymphoma; wound healing; human; chromosome 15 q22.3-23.

OS Homo sapiens.

XX WO200029435-A1.

XX 25-MAY-2000.

XX 27-OCT-1999; 99WO-US025031.

XX 28-OCT-1998; 98US-0105971P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;

PI Greene JW;

XX WPI; 2000-387742/33.

XX Isolated nucleic acid molecules encoding human secreted proteins are used

PT for the prevention, amelioration and treatment of autoimmune,

PT inflammatory, hyperproliferative and cardiovascular disorders, cancer,

PT wounds, and infectious diseases.

XX Claim 1; Page 678-682; 803pp; English.

XX The present invention relates to 12 secreted human proteins and the

CC nucleotide sequences encoding them. The polynucleotide sequences given in

CC AAA80606-A80623 encode the 12 secreted protein sequences given in

CC AAB25576-B25593. The human secreted proteins have various activities

CC dependent on the tissues in which they are expressed. Examples of the

CC activities of the proteins include: immunosuppressant; anti-inflammatory;

CC antiarthritic; antirheumatic; dermatological; antiproliferative;

CC antiarteriosclerotic; anticancer; vulnery; antiviral; antibacterial;

CC and antifungal activity. The proteins, polypeptides, agonists and

CC antagonists may be used to treat prevent and/or diagnose various disease,

CC disorders and conditions examples of which include: immune disorders e.g.

CC Addison's disease, rheumatoid arthritis, dermatitis, and multiple

CC sclerosis; inflammatory disorders e.g. inflammatory bowel disease,

CC Crohn's disease and nephritis; hyperproliferative disorders such as

CC paraproteinemias and purpura; cardiovascular disorders e.g. coronary

CC arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The

CC proteins and polynucleotide sequences may also be used in wound healing

CC and the treatment of infectious diseases. The human secreted protein gene

CC #7 and protein sequences are represented in sequences AAA80612 and

CC AAB25582. Secreted protein gene #7 is located at position chromosome 15

CC q22.3-23. Sequences AAA80652-A80661 represent genes which are related to

CC the secreted protein gene#7

XX SQ Sequence 1034 AA;

Query Match 68.2%; Score 810; DB 3; Length 1034;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDLPRGLVVAWALSLWPGFTDTFNMDTRKPRVPGSRTPAFGYTVQOHDISGNKWLWVGA 60

Db 1 MDLPRGLVVAWALSLWPGFTDTFNMDTRKPRVPGSRTPAFGYTVQOHDISGNKWLWVGA 60

QY 61 PLETMGYOKTGDVYKCPVIHGNCTKLNLRGVTLSNVSRKDNMRGLSLATLPKDNSFLA 120

DB 61 PLETMGYOKTGDVYKCPVIHGNCTKLNLRGVTLSNVSRKDNMRGLSLATLPKDNSFLA 120

QY 121 CSPLMSEHGSSYYTTGMCNRVNSNFRFSKTVAPALQRCQTMQMDIVIVLQNSNYSYPWVE 180

DB 121 CSPLMSEHGSSYYTTGMCNRVNSNFRFSKTVAPALQRCQTMQMDIVIVLQNSNYSYPWVE 180

QY 181 VOHFLINILKKFYIGPGQIQGVVQYGVGDDVHEPHLNDYRSVKDWEAASHTEORGGTET 240

DB 181 VOHFLINILKKFYIGPGQIQGVVQYGVGDDVHEPHLNDYRSVKDWEAASHTEORGGTET 240

QY 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGHSDSPDLKVIQOQSRDNVTRYAVAVL 300

DB 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGHSDSPDLKVIQOQSRDNVTRYAVAVL 300

QY 301 GYNRRGINPETFLNEIKYIASDDDDKHFFNVTEAALKIDVADGDRIFSLGNTKNET 360

DB 301 GYNRRGINPETFLNEIKYIASDDDDKHFFNVTEAALKIDVADGDRIFSLGNTKNET 360

QY 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDMNGAVLKETSAKVILPRLSYLKEPPEELKN 420

DB 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDMNGAVLKETSAKVILPRLSYLKEPPEELKN 420

QY 421 HGAYLGYTVTSVWSRQGRVYVAGAPRNHTGKVLFTMNNRSITIHQMRGQOIGSYF 480

DB 421 HGAYLGYTVTSVWSRQGRVYVAGAPRNHTGKVLFTMNNRSITIHQMRGQOIGSYF 480

QY 481 GSEITSVDIDGQVTDVLLVAGAPMYFNEGRGRGVYVYELRQNRVYVNGTLKDSHSYQNA 540

DB 481 GSEITSVDIDGQVTDVLLVAGAPMYFNEGRGRGVYVYELRQNRVYVNGTLKDSHSYQNA 540

QY 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFRGSILKTPKQRTASELATG 600

DB 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFRGSILKTPKQRTASELATG 600

QY 601 LQYFCCSHGQLDNEGLIDLAVGALGNVILMSRPVVQINASLHFPSPKINIPIHRDCK 660

DB 601 LQYFCCSHGQLDNEGLIDLAVGALGNVILMSRPVVQINASLHFPSPKINIPIHRDCK 660

QY 661 RSGRDATCLAFCTPTPIFLAPHQFTTVTGIRYNATMDERRYTTPRAHLDEGGDRFTNRAV 720

DB 661 RSGRDATCLAFCTPTPIFLAPHQFTTVTGIRYNATMDERRYTTPRAHLDEGGDRFTNRAV 720

QY 721 LLSSQBELCERINFHVLDTADYVKVPTSVESYLEDDPHGPMLDGMPDGLLVSVYFVWNG 780

DB 721 LLSSQBELCERINFHVLDTADYVKVPTSVESYLEDDPHGPMLDGMPDGLLVSVYFVWNG 780

QY 781 CNEDEHCVPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTILSPDTTFFIIBSTQRVAV 840

DB 781 CNEDEHCVPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTILSPDTTFFIIBSTQRVAV 840

QY 841 EATLENRGENAYSTVLNITSQANLQFASLIQKEDSDGSIQCVNEERRLQKQVCNYSYPPF 900

DB 841 EATLENRGENAYSTVLNITSQANLQFASLIQKEDSDGSIQCVNEERRLQKQVCNYSYPPF 900

QY 901 RAKAKVAFRLD 911

DB 901 RAKAKVAFRLD 911

RESULT 14

ADA27062

ID ADA27062 standard; protein; 1034 AA.

XX ADA27062;

XX AC ADA27062;

DT 20-NOV-2003 (first entry)

XX Human novel secreted protein from cDNA HOHB69 #2.

XX cytostatic; antiinflammatory; immunomodulator; neuroprotective;

KW

hemostatic; gene therapy; cancer; inflammation; immune disorder; neurological disorder; blood clotting disorder; food additive; preservative; human; secreted protein.

Homo sapiens.

US2003055231-A1.

20-MAR-2003.

29-OCT-2001; 2001US-00984130.

28-OCT-1998; 98US-0105971P.

27-OCT-1999; 99WO-US025031.

19-APR-2000; 2000US-0198407P.

30-OCT-2000; 2000US-0243792P.

18-APR-2001; 2001US-00836353.

(NIJJ/) NI J.

(YOUN/) YOUNG P B.

(KENN/) KENNY J J.

(OLSE/) OLSEN H S.

(MOOR/) MOORE P A.

(WEIY/) WEI Y.

(GREE/) GREENE J M.

(RUBE/) RUBEN S M.

(LIUD/) LIU D.

(CROC/) CROCKER P R.

Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM; Ruben SM, Liu D, Crocker PR; WPI: 2003-567103/53. N-PSDB; ADA27044.

New human secreted nucleic acid molecules and polypeptides, useful for preventing, treating, or ameliorating a medical condition, such as cancer, inflammation, immune disorders, neurological and blood clotting disorders.

Claim 11; Page 302-305; 454pp; English.

The invention relates to an isolated nucleic molecule that is at least 95% identical to 18 human cDNA sequences representing 12 novel genes encoding secreted proteins or a polynucleotide fragment of the cDNA sequence contained in American Type Culture Collection (ATCC) deposit No. defined in the specification, its species homologue, a variant or allelic variant of the polynucleotide having a polynucleotide capable of hybridising under conditions the polynucleotide, where the polynucleotide does not hybridise under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A or T residues. Also included are recombinant vectors, host cells (for producing the polypeptide), the secreted polypeptide (comprising a sequence that is at least 95% identical to a polypeptide fragment, domain, epitope, full-length protein, variant, allelic variant or species homologue), antibodies that specifically bind to the polypeptides, diagnosing, treating, preventing or ameliorating a medical condition by administering the polynucleotide or the polypeptide, the gene corresponding to the cDNA sequence and identifying an activity in a biological assay (by expressing the cDNA sequence in a cell, isolating the supernatant, and detecting an activity in a biological assay and identifying the protein in the supernatant having the activity). The polypeptides, nucleic acids and antibodies are useful for diagnosing a pathological condition or a susceptibility to a pathological condition, for preventing, treating, or ameliorating a medical condition, such as cancer, inflammation and other immune disorders, neurological and blood clotting disorders (many examples are given in the specification). The nucleic acids are also useful for chromosome identification, radiation hybrid mapping or long-range restriction mapping. The polypeptides and antibodies are useful for providing immunological probes for differential identification of the tissues immunohistochemistry assays. The polypeptide, polynucleotide, agonist or antagonist may also be used as a food additive or preservative to increase or decrease storage capabilities, fat content or other

CC nutritional components. The present is a secreted protein of the  
CC invention.  
XX  
XX  
SQ Sequence 1034 AA;  
Query Match 68.2%; Score 310; DB 6; Length 1034;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MDLPRLGVVAWALSMPGFTDTFNMDTRKPRVPGSRTAFPGYTVQGHDSGNKWLVGA 60  
Db 1 MDLPRLGVVAWALSMPGFTDTFNMDTRKPRVPGSRTAFPGYTVQGHDSGNKWLVGA 60  
Qy 61 PLETNGYQKTGDYKCPVHGNCTKLNIGVTLNVSERKONRGLSLATNPKDNSFLA 120  
Db 61 PLETNGYQKTGDYKCPVHGNCTKLNIGVTLNVSERKONRGLSLATNPKDNSFLA 120  
Qy 121 CSPLWSHECGSSYYTGMCSRVSNSRFSKTVAPALQRCQTYMDIVILVDSNSIYEWVE 180  
Db 121 CSPLWSHECGSSYYTGMCSRVSNSRFSKTVAPALQRCQTYMDIVILVDSNSIYEWVE 180  
Qy 181 VOHFLINILKIFYIGGQIQGVVQGVGEDVHFLNDYRSVKDVVAAASHIQRGGTET 240  
Db 181 VOHFLINILKIFYIGGQIQGVVQGVGEDVHFLNDYRSVKDVVAAASHIQRGGTET 240  
Qy 241 RTAFGIEFARSEAFQKGRKGAKKMWIIVTDGSHSDPDLKVIQSERDNVTRYAVAVL 300  
Db 241 RTAFGIEFARSEAFQKGRKGAKKMWIIVTDGSHSDPDLKVIQSERDNVTRYAVAVL 300  
Qy 301 GYNNRGINPETFLNEIKYIASDPDDKHFNVTDEAALKDIVDALGRIIFSLEGTNKET 360  
Db 301 GYNNRGINPETFLNEIKYIASDPDDKHFNVTDEAALKDIVDALGRIIFSLEGTNKET 360  
Qy 361 SFGLEMSQTCFSSHVEDGVLLGAVGAYDMNGAVLAKTSAGKVIPLRESYLKPEPELKN 420  
Db 361 SFGLEMSQTCFSSHVEDGVLLGAVGAYDMNGAVLAKTSAGKVIPLRESYLKPEPELKN 420  
Qy 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTGKVIPLFMNRSLSLTHQAMRQGIISYF 480  
Db 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTGKVIPLFMNRSLSLTHQAMRQGIISYF 480  
Qy 481 GSITSVDIDGDTVDVLLGAPMYNEGRGKVVYVELRQNFVNGTLKSHSYQNA 540  
Db 481 GSITSVDIDGDTVDVLLGAPMYNEGRGKVVYVELRQNFVNGTLKSHSYQNA 540  
Qy 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNEAGAIYIFHGFSGSILKTPKQRIITASELATG 600  
Db 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNEAGAIYIFHGFSGSILKTPKQRIITASELATG 600  
Qy 601 LQVFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPVQVNASLHPEPSKINIFPHROCK 660  
Db 601 LQVFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPVQVNASLHPEPSKINIFPHROCK 660  
Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTPRAHLDGEGDFTNRVAV 720  
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTPRAHLDGEGDFTNRVAV 720  
Qy 721 LLSGGQELCERINPHVLDYADYKVPVTFVSVEYSLEDHPGMLDDGHPPTILRVSPVFWNG 780  
Db 721 LLSGGQELCERINPHVLDYADYKVPVTFVSVEYSLEDHPGMLDDGHPPTILRVSPVFWNG 780  
Qy 781 CNEDEHCVDPDLVADARSDLPTAMEYCORVLKPAQDCSATLSFDFTTVFIIESTRQRAVAV 840  
Db 781 CNEDEHCVDPDLVADARSDLPTAMEYCORVLKPAQDCSATLSFDFTTVFIIESTRQRAVAV 840  
Qy 841 EATLENREGENAYSTVLNISQSANIQFASLIQKEDSDGSIKEDNERSLQKQCNVSPYFF 900  
Db 841 EATLENREGENAYSTVLNISQSANIQFASLIQKEDSDGSIKEDNERSLQKQCNVSPYFF 900  
Qy 901 RAKAKVAFRLD 911  
Db 901 RAKAKVAFRLD 911

RESULT 15  
 ADE86592  
 ID ADE86592 standard; protein; 1034 AA.  
 AC ADE86592;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Novel human secreted protein #15.  
 XX  
 KW human; secreted protein; cancer; liver disorder; hepatitis;  
 KW neural disorder; Alzheimer's disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003129685-A1.  
 XX  
 PD 10-JUL-2003.  
 XX  
 PF 18-APR-2001; 2001US-00836353.  
 XX  
 PR 28-OCT-1998; 98US-0105971P.  
 PR 27-OCT-1999; 99WO-US025031.  
 PR 19-APR-2000; 2000US-0198407P.  
 XX  
 PA (NIJJ/) NI J.  
 PA (YOUN/) YOUNG P E.  
 PA (KENN/) KENNY J J.  
 PA (OLSE/) OLSEN H S.  
 PA (MOOR/) MOORE P A.  
 PA (WEIY/) WEI Y.  
 PA (GREG/) GREENE J M.  
 PA (RUBE/) RUBEN S M.  
 XX  
 FI Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;  
 PI Ruben SM;  
 XX  
 DR WPI; 2004-020335/02.  
 DR N-PSDB; ADE86574.  
 XX  
 PT New nucleic acid molecule, useful for preparing a medicament for  
 PT preventing, treating or ameliorating a medical condition e.g. cancer,  
 PT liver disorders or neural disorders.  
 XX  
 PS Claim 11; SEQ ID NO 43; 380pp; English.  
 CC  
 CC The invention relates to an isolated nucleic acid sequence, or its  
 CC allelic variant, a fragment of the cDNA sequence, or its fragment,  
 CC domain, epitope or species homologue. The nucleic acid is useful for  
 CC preparing a medicament for preventing, treating or ameliorating a medical  
 CC condition e.g., cancer, liver disorders such as hepatitis or neural  
 CC disorders such as Alzheimer's disease. The present sequence represents  
 CC the amino acid sequence of a novel human secreted protein.  
 XX  
 SQ Sequence 1034 AA;  
 Query Match 68.2%; Score 810; DB 8; Length 1034;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 MDLPRGLVAVALSFWPGFTDFTNMDTRKPRVPGSTAFPGYTVQOHDISGNKWLAVGA 60  
 Db 1 MDLPRGLVAVALSFWPGFTDFTNMDTRKPRVPGSTAFPGYTVQOHDISGNKWLAVGA 60  
 Qy 61 PLETNGYQKTGDVYKCPVHGNCTKLNIGRVTLSNVSRKDNMRGLGLSLATNPKNLSFLA 120  
 Db 61 PLETNGYQKTGDVYKCPVHGNCTKLNIGRVTLSNVSRKDNMRGLGLSLATNPKNLSFLA 120  
 Qy 121 CSPLWSEHCSSYYTTCMCSRVSNSFRFSKTVPALQRCQTYMDIVIVLDGNSIYPWVE 180  
 Db 121 CSPLWSEHCSSYYTTCMCSRVSNSFRFSKTVPALQRCQTYMDIVIVLDGNSIYPWVE 180

QY 181 VQHFLLINILKKFYIGFGQIQGVVQYGEDVWHEFLNDYRSVKDWEAAASHIEORGGTET 240  
 DB 181 VQHFLLINILKKFYIGFGQIQGVVQYGEDVWHEFLNDYRSVKDWEAAASHIEORGGTET 240  
 QY 241 RTAFGIEFARSEAFQGRKGAKKVMIVITDSESHSDPLEKVIQOSENDRNTRAVAVL 300  
 DB 241 RTAFGIEFARSEAFQGRKGAKKVMIVITDSESHSDPLEKVIQOSENDRNTRAVAVL 300  
 QY 301 GYNNRRGINPETFLNEIKYIASDPDCKHFFNVYTDAAALKDIVDALGDRIFSLGNTKNKET 360  
 DB 301 GYNNRRGINPETFLNEIKYIASDPDCKHFFNVYTDAAALKDIVDALGDRIFSLGNTKNKET 360  
 QY 361 SFGLENSQTFSSSHVDEGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPBELKN 420  
 DB 361 SFGLENSQTFSSSHVDEGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPBELKN 420  
 QY 421 HGAYLGYTVTSVVSRRQGRVYVAGAPRNFHTCKVILFTMHNNRSLTIHOAMRGQIQISYF 480  
 DB 421 HGAYLGYTVTSVVSRRQGRVYVAGAPRNFHTCKVILFTMHNNRSLTIHOAMRGQIQISYF 480  
 QY 481 GSEITSDVIDGQVTDVLLVGAPMFNREGREGKVVYVELRQNRFPYNGTLKDSHSYQNA 540  
 DB 481 GSEITSDVIDGQVTDVLLVGAPMFNREGREGKVVYVELRQNRFPYNGTLKDSHSYQNA 540  
 QY 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGALVIFHGFGRGSLKTPKORITASELATG 600  
 DB 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGALVIFHGFGRGSLKTPKORITASELATG 600  
 QY 601 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPVVQINASLHFEPESKINIHRDCK 660  
 DB 601 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPVVQINASLHFEPESKINIHRDCK 660  
 QY 661 RSGRDATCLAAFLCETPIFLAPHFQTTVIGIRYNATMDERRYTPRAHLDEGGDRFTNRAV 720  
 DB 661 RSGRDATCLAAFLCETPIFLAPHFQTTVIGIRYNATMDERRYTPRAHLDEGGDRFTNRAV 720  
 QY 721 LLSGQBELCERINFHVLDTADYVKPVTFSVEYSLEDDPHGPMDDGWPPTLTVSVVPFWNG 780  
 DB 721 LLSGQBELCERINFHVLDTADYVKPVTFSVEYSLEDDPHGPMDDGWPPTLTVSVVPFWNG 780  
 QY 781 CNEDEHCYVDLVDARSDLPTAMEYCORVLRKPAQCSAYTSLFDTTVFIIESTQRVAV 840  
 DB 781 CNEDEHCYVDLVDARSDLPTAMEYCORVLRKPAQCSAYTSLFDTTVFIIESTQRVAV 840  
 QY 841 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGSECYNEERLQKQVCNVSYPFF 900  
 DB 841 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGSECYNEERLQKQVCNVSYPFF 900  
 QY 901 RAKAKVAFRLD 911  
 DB 901 RAKAKVAFRLD 911

RESULT 16

ABR58365  
 ID ABR58365 standard; protein; 1120 AA.  
 XX  
 AC ABR58365;  
 XX  
 DT 07-JUL-2003 (first entry)  
 XX  
 DE Human NOV2b.  
 XX  
 KW Human; NOV; antidiabetic; anorectic; antibacterial; virucide;  
 KW immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia;  
 KW antiparkinsonian; antilipaeamic; gene therapy; metabolic disorder;  
 KW diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;  
 KW neurodegenerative disorder; Alzheimer's disease; immune disorder;  
 KW haematopoietic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003029423-A2.

XX 10-APR-2003.  
 XX 02-OCT-2002; 2002WO-US031358.  
 XX 02-OCT-2001; 2001US-0326483P.  
 XX 05-OCT-2001; 2001US-0327342P.  
 XX 09-OCT-2001; 2001US-0327517P.  
 XX 09-OCT-2001; 2001US-0328029P.  
 XX 09-OCT-2001; 2001US-0328044P.  
 XX 09-OCT-2001; 2001US-0328056P.  
 XX 12-OCT-2001; 2001US-0328849P.  
 XX 15-OCT-2001; 2001US-0329414P.  
 XX 17-OCT-2001; 2001US-0330142P.  
 XX 22-OCT-2001; 2001US-0341058P.  
 XX 24-OCT-2001; 2001US-0339266P.  
 XX 24-OCT-2001; 2001US-0343629P.  
 XX 29-OCT-2001; 2001US-0349575P.  
 XX 01-NOV-2001; 2001US-0346357P.  
 XX 12-APR-2002; 2002US-0371972P.  
 XX 12-APR-2002; 2002US-0371980P.  
 XX 17-APR-2002; 2002US-0373261P.  
 XX 19-APR-2002; 2002US-0373805P.  
 XX 23-APR-2002; 2002US-0374738P.  
 XX 16-MAY-2002; 2002US-0381101P.  
 XX 17-MAY-2002; 2002US-0381635P.  
 XX 29-MAY-2002; 2002US-0383830P.  
 XX 01-OCT-2002; 2002US-00282839.  
 XX (CURA-) CURAGEN CORP.  
 XX Alsbrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;  
 XX Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;  
 XX Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;  
 XX Rothenberg ME, Shimkets RA, Smithson G, Spytek KA, Taupier RJ;  
 XX Vernet CM, Voss EZ, Zernhusen BD, Zhong M;  
 XX WPI; 2003-381625/36.  
 XX N-PSDB; ACC72077.  
 XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or  
 XX treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or  
 XX dyslipidemia, and in chromosome mapping, tissue typing or  
 XX pharmacogenomics.  
 XX Claim 1; Page 107; 487pp; English.  
 XX The present invention relates to novel human NOV proteins and their  
 XX coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV  
 XX proteins are useful in manufacturing a medicament for treating a syndrome  
 XX associated with a human disease. The NOV proteins and coding sequences  
 XX may be used to diagnose, treat or prevent metabolic disorders such as  
 XX diabetes or obesity, infections, cachexia, cancer, neurodegenerative  
 XX disorders such as Alzheimer's disease or Parkinson's disease, immune  
 XX disorders, haematopoietic disorders and various dyslipidaemias  
 XX Sequence 1120 AA;  
 XX Query Match 63.4%; Score 753; DB 6; Length 1120;  
 XX Best Local Similarity 100.0%; Pred. No. 0;  
 XX Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX 159 CQYMDIVIVLDGNSIYPVWEVQHFILNLIKFPVIGCQIQGVGVQGVGVVHFLND 218  
 XX |||||  
 XX 90 CQYMDIVIVLDGNSIYPVWEVQHFILNLIKFPVIGCQIQGVGVQGVGVVHFLND 149  
 XX |||||  
 XX 219 YRSVKDVEAASHIQRGGTERTAFGIEFARSEAFQGRGAKKVMIVITDGHSDSP 278  
 XX |||||  
 XX 150 YRSVKDVEAASHIQRGGTERTAFGIEFARSEAFQGRGAKKVMIVITDGHSDSP 209  
 XX |||||  
 XX 279 DLEKVIQSERDNVTRYAVAVLGYNNRGINPETFNLNKKIYIASDPDDKHFNVTEAAL 338  
 XX |||||  
 XX 210 DLEKVIQSERDNVTRYAVAVLGYNNRGINPETFNLNKKIYIASDPDDKHFNVTEAAL 269  
 XX |||||

QY 339 KDIVDALGDRIFFSLEGTNRKNETSFGLSEMSQTFSSHVVEDGVLLGAYGAVMNGAVLKET 398  
 DB 270 KDIVDALGDRIFFSLEGTNRKNETSFGLSEMSQTFSSHVVEDGVLLGAYGAVMNGAVLKET 329  
 QY 399 SAGKVIPLRESYLKEPPEELKNHGAYLGYTVTSVSSROGKYVYAGAPRPHNHTKVLFT 458  
 DB 330 SAGKVIPLRESYLKEPPEELKNHGAYLGYTVTSVSSROGKYVYAGAPRPHNHTKVLFT 389  
 QY 459 MHNRSITIHQAMEGQQIGSYFGSEITSVDIDGQVTDVLLVGAPMTFNEGRERGVVYV 518  
 DB 390 MHNRSITIHQAMEGQQIGSYFGSEITSVDIDGQVTDVLLVGAPMTFNEGRERGVVYV 449  
 QY 519 ELQRNRFVYNGTLKDSHYQNAFPGSSIASVRDLNODSYNDVYVGALEDNHAGAIYF 578  
 DB 450 ELQRNRFVYNGTLKDSHYQNAFPGSSIASVRDLNODSYNDVYVGALEDNHAGAIYF 509  
 QY 579 GFRGSILKTPKQIRITASELATGLQYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPV 638  
 DB 510 GFRGSILKTPKQIRITASELATGLQYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPV 569  
 QY 639 VQINASHFEPFSKINIFHRDCKSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMD 698  
 DB 570 VQINASHFEPFSKINIFHRDCKSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMD 629  
 QY 699 ERYTTPRAHLDGGRFTNRAVLLSSQCELCERINFHVLDTADYVKVPTSVESYLEDDP 758  
 DB 630 ERYTTPRAHLDGGRFTNRAVLLSSQCELCERINFHVLDTADYVKVPTSVESYLEDDP 689  
 QY 759 HGPMLDDGWPFTTLRVSVFPWNGCNEDEHCVPLDVLARSOLPTAMEYQVRVLRPAQDCS 818  
 DB 690 HGPMLDDGWPFTTLRVSVFPWNGCNEDEHCVPLDVLARSOLPTAMEYQVRVLRPAQDCS 749  
 QY 819 AYTLSEFTTTVFILTESTRQAVATLENRGENAVSTVLNISQSANLOFASLIQKEDSDGS 878  
 DB 750 AYTLSEFTTTVFILTESTRQAVATLENRGENAVSTVLNISQSANLOFASLIQKEDSDGS 809  
 QY 879 IECVNEERRLQKQVCNVSYPPFFRAKAVAFELD 911  
 DB 810 IECVNEERRLQKQVCNVSYPPFFRAKAVAFELD 842  
 RESULT 17  
 AAU19663  
 ID AAU19663 standard; protein; 707 AA.  
 XX AC AAU19663;  
 XX DT 04-DEC-2001 (first entry)  
 XX DE Human novel extracellular matrix protein, Seq ID No 313.  
 XX KW Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;  
 KW antianemic; antirheumatic; antisclerotic; cardiac; vascular;  
 KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;  
 KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;  
 KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;  
 KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;  
 KW Sezary syndrome; Gaucher's disease; neurological diseases;  
 KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
 KW cardiac arrest; tachycardia; angina; infection; corneal infections;  
 KW wound healing; immunogen; gene therapy; antisense; food additive.  
 XX OS Homo sapiens.  
 XX FN WO200155368-A1.  
 XX PD 02-AUG-2001.  
 XX PF 17-JAN-2001; 2001WO-US001348.  
 XX PR 31-JAN-2000; 2000US-0179065P.  
 XX PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 01-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235634P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465572/50.  
XX N-PSDB; AAS31234.  
XX  
XX Nucleic acid molecules encoding human secreted extracellular matrix  
XX proteins, used in preventing, treating or ameliorating a disorder, e.g.  
XX Alzheimer's and Parkinson's diseases and cancers.

PS	Claim 11; SEQ ID NO 313; 577bp; English.	
XX	The invention relates to isolated nucleic acid molecules encoding novel	
CC	human secreted extracellular matrix proteins (SPs). The polynucleotides	
CC	and proteins are used to prevent, treat a medical condition in e.g.	
CC	humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. For	
CC	example, disorders associated with decreased expression of SPs. The SP	
CC	polynucleotide or a vector expressing them may be administered to treat	
CC	diseases by gene therapy. Antisense molecules may be administered to down	
CC	regulate expression of SPs by binding with the cells own genes and	
CC	preventing their expression. The polynucleotides may also be used as DNA	
CC	probes in diagnostic assays. The SPs may also be used as antigens to	
CC	produce antibodies and to identify modulators (agonists and antagonists)	
CC	of the SPs. The anti-(SP) antibodies and antagonists may also be used to	
CC	down regulate expression and activity of SP and as diagnostic agents for	
CC	detecting the presence of SPs in samples. The disorders include for	
CC	example: immune/autoimmune diseases (e.g. HIV (human immunodeficiency	
CC	virus) infections, anaemia, rheumatoid arthritis and multiple sclerosis),	
CC	cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of	
CC	the breast or liver, Sezary syndrome and Gaucher's disease), neurological	
CC	diseases (e.g. Alzheimer's disease, Parkinson's disease) cardio-	
CC	/cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina),	
CC	infections caused by bacteria, viruses and fungi and ocular disorders	
CC	(e.g. corneal infections). Other uses include wound healing, maintenance	
CC	of organs before transplantation, support of cell culture of primary	
Query Match 41.2%; Score 489; DB 4; Length 707;		
Best Local Similarity 99.7%; Pred. No. 0;		
Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Qy	1 MDLPRGLVVAWALSMPGFTDFTNMDTRKPRVPGSRVTAFFGVTYQQHDSGNKWLVA 60	
Db	17 MDLPRGLVVAWALSMPGFTDFTNMDTRKPRVPGSRVTAFFGVTYQQHDSGNKWLVA 76	
Qy	61 PLETNGYQKTDGVYKCPVTHGNCIKLGRVTLNVSERKONRGLSLATNPKNSFLA 120	
Db	77 PLETNGYQKTDGVYKCPVTHGNCIKLGRVTLNVSERKONRGLSLATNPKNSFLA 136	
Qy	121 CSPLSHSCGSSYTTTGMCSRNSNFRSKTVAPALQRCQTYMDIVIVLDGNSLYPWVE 180	
Db	137 CSPLSHSCGSSYTTTGMCSRNSNFRSKTVAPALQRCQTYMDIVIVLDGNSLYPWVE 196	
Qy	181 VQHFLNLIKFFYIGPQIQGVQVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 240	
Db	197 VQHFLNLIKFFYIGPQIQGVQVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 256	
Qy	241 RTAFGEFARSEAFQGGKRGAKKVMIVITDGHSDSPLEKVIQSERDNTVRVAVL 300	
Db	257 RTAFGEFARSEAFQGGKRGAKKVMIVITDGHSDSPLEKVIQSERDNTVRVAVL 316	
Qy	301 GYNNRGINPETFLEIKYIASDDDKHFNVTDEAALKDIDVDAALGDRIFSLGNTKNET 360	
Db	317 GYNNRGINPETFLEIKYIASDDDKHFNVTDEAALKDIDVDAALGDRIFSLGNTKNET 376	
Qy	361 SFGLEMSQTFSSHVVEDGVLLGAVCAVWNGAVLKETSGAKVILPRESYLKEFFPELQN 420	
Db	377 SFGLEMSQTFSSHVVEDGVLLGAVCAVWNGAVLKETSGAKVILPRESYLKEFFPELQN 436	
Qy	421 HGAYLGYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMHNNRSLTIHQAMRGQQIGSYF 480	
Db	437 HGAYLGYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMHNNRSLTIHQAMRGQQIGSYF 496	
Qy	481 GSITSDVDIGDGVTVLLVGAIPWYFNEGRERKGVVYVELRQNRFYNGTLKDSHXYQNA 540	
Db	497 GSITSDVDIGDGVTVLLVGAIPWYFNEGRERKGVVYVELRQNRFYNGTLKDSHXYQNA 556	
Qy	541 RFGSSIASVRDLNODSNDVWVCAPLEDNHAGAIYIFPHGRSILKTPKQRTASBLATG 600	
Db	557 RFGSSIASVRDLNODSNDVWVCAPLEDNHAGAIYIFPHGRSILKTPKQRTASBLATG 616	
Qy	601 LQYFGCSIHQQLDNLNEDGLDLAVGALGNVILWSPVQVQINASLHFEPSKINIFHRDCK 660	
Db	617 LQYFGCSIHQQLDNLNEDGLDLAVGALGNVILWSPVQVQINASLHFEPSKINIFHRDCK 676	

Qy	661 RSGRDATCAAFLCFTFIPLAPHFQTTTNGI 691	
Db	677 RSGRDATCAAFLCFTFIPLAPHFQTTTNGI 707	
RESULT 18		
ABP47883		
ID	ABP47883 standard; protein; 707 AA.	
XX	ABP47883;	
XX	23-AUG-2002 (first entry)	
DT	Human polypeptide SEQ ID NO 313.	
DE	Human; nontropic; neuroprotective; cytostatic; dermatological; virucide;	
XX	immunosuppressive; antineoplastic; anti-HIV; antibacterial; vulnary;	
XX	antiparkinsonian; antischizoid; antianemic; antidiabetic; cancer;	
XX	antiallergic; hepatotropic; cerebroprotective; antiinflammatory;	
XX	antiparasitic; cardiac; antitumor; anticonvulsant; antifungal;	
XX	neurological disease; infection; nephrotropic; gene therapy; vaccine.	
OS	Homo sapiens.	
PN	US2002042386-A1.	
XX	11-APR-2002.	
PF	17-JAN-2001; 2001US-00764870.	
PR	31-JAN-2000; 2000US-0179065P.	
PR	04-FEB-2000; 2000US-0180628P.	
PR	28-JUN-2000; 2000US-0214886P.	
PR	07-JUL-2000; 2000US-0216647P.	
PR	07-JUL-2000; 2000US-0216880P.	
PR	11-JUL-2000; 2000US-0217487P.	
PR	14-JUL-2000; 2000US-0218290P.	
PR	26-JUL-2000; 2000US-0220963P.	
PR	14-AUG-2000; 2000US-0224518P.	
PR	14-AUG-2000; 2000US-0224519P.	
PR	14-AUG-2000; 2000US-0225267P.	
PR	14-AUG-2000; 2000US-0225268P.	
PR	14-AUG-2000; 2000US-0225447P.	
PR	14-AUG-2000; 2000US-0225757P.	
PR	14-AUG-2000; 2000US-0225758P.	
PR	22-AUG-2000; 2000US-0226868P.	
PR	30-AUG-2000; 2000US-0228924P.	
PR	01-SEP-2000; 2000US-0229287P.	
PR	01-SEP-2000; 2000US-0229343P.	
PR	01-SEP-2000; 2000US-0229344P.	
PR	01-SEP-2000; 2000US-0229345P.	
PR	05-SEP-2000; 2000US-0229509P.	
PR	05-SEP-2000; 2000US-0229513P.	
PR	08-SEP-2000; 2000US-0231413P.	
PR	21-SEP-2000; 2000US-0234223P.	
PR	21-SEP-2000; 2000US-0234274P.	
PR	25-SEP-2000; 2000US-0234997P.	
PR	27-SEP-2000; 2000US-0235834P.	
PR	29-SEP-2000; 2000US-0236327P.	
PR	29-SEP-2000; 2000US-0236367P.	
PR	29-SEP-2000; 2000US-0236368P.	
PR	29-SEP-2000; 2000US-0236369P.	
PR	02-OCT-2000; 2000US-0236370P.	
PR	02-OCT-2000; 2000US-0236802P.	
PR	02-OCT-2000; 2000US-0237037P.	
PR	02-OCT-2000; 2000US-0237038P.	
PR	02-OCT-2000; 2000US-0237039P.	
PR	02-OCT-2000; 2000US-0237040P.	

PR 13-OCT-2000; 2000US-0239935P.  
 PR 20-OCT-2000; 2000US-0240950P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PA (ROSE/) ROSEN C A.  
 PA (RUBE/) RUBEN S M.  
 PA (BARA/) BARASH S C.  
 XX  
 PI Rosen CA, Ruben SM, Barash SC;  
 XX  
 DR W21; 2002-470713/50.  
 DR N-PSDB; ABQ66558.  
 XX  
 PT New nucleic acid encoding human proteins, useful for diagnosis, treatment  
 PT and prevention of e.g. osteoporosis, also related polypeptides and  
 PT antibodies.  
 XX  
 PS Claim 11; SEQ ID NO 313; 235pp + Sequence Listing; English.  
 XX  
 CC The invention relates to novel genes (ABQ66521-ABQ66785) and proteins  
 CC (ASP47846-ASP48110) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital, (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from USPTO at seqdata.uspto.gov/sequence.html?DocID=999909764870  
 XX  
 SQ Sequence 707 AA;

Query Match 41.2%; Score 489; DB 5; Length 707;  
 Best Local Similarity 99.7%; Pred No. 0;  
 Matches 589; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MDLPRGLVAVWALSMPGFTDTNMDTRKPRVPGSRTAPFGYTVQVQHDISGNKMLVWGA 60  
 Db 17 MDLPRGLVAVWALSMPGFTDTNMDTRKPRVPGSRTAPFGYTVQVQHDISGNKMLVWGA 76  
 Qy 61 PLENTGYOKTGDYKCPVIEHGNCTKLNLRVTLSNVYSERKDNRLGLSLATNPKNLSFLA 120  
 Db 77 PLENTGYOKTGDYKCPVIEHGNCTKLNLRVTLSNVYSERKDNRLGLSLATNPKNLSFLA 136  
 Qy 121 CSPLWSHCEGSSYFTTGMCSRVSNSFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPPWVE 180  
 Db 137 CSPLWSHCEGSSYFTTGMCSRVSNSFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPPWVE 196  
 Qy 181 VQFLINILKFFVIGPQIQGVVQYGEDVWHFPHLNDVRSVKDVAASHIEQRGTTT 240  
 Db 197 VQFLINILKFFVIGPQIQGVVQYGEDVWHFPHLNDVRSVKDVAASHIEQRGTTT 256  
 Qy 241 RTAFGIEFARSEAFQGRGKAGKVMIVITDGSHPDLEKVIQSERDNTVRYAVVL 300  
 Db 257 RTAFGIEFARSEAFQGRGKAGKVMIVITDGSHPDLEKVIQSERDNTVRYAVVL 316  
 Qy 301 GYNNRGINPTELNEIKYIASDPDDKHFFNVTDAAKLDIVDALGDRIFSLGTTKNET 360  
 Db 317 GYNNRGINPTELNEIKYIASDPDDKHFFNVTDAAKLDIVDALGDRIFSLGTTKNET 376

Qy 361 SFGLEMSOTGFSHSHVVDGVLGAVGYDNGAVLKVTSAGKVIPLRESYLKEFPEELKN 420  
 Db 377 SFGLEMSOTGFSHSHVVDGVLGAVGYDNGAVLKVTSAGKVIPLRESYLKEFPEELKN 436  
 Qy 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSILTIHQAMRGQOIGSYF 480  
 Db 437 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSILTIHQAMRGQOIGSYF 496  
 Qy 481 GSEITSVDIDGVDVLLVAGAPYFNEGRERGVYVYELRQNRVFNVTGLKXDSHSYQNA 540  
 Db 497 GSEITSVDIDGVDVLLVAGAPYFNEGRERGVYVYELRQNRVFNVTGLKXDSHSYQNA 556  
 Qy 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGFRGSILKTKPQRTASELATG 600  
 Db 557 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGFRGSILKTKPQRTASELATG 616  
 Qy 601 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILWSPVVOINASLHFFPSKINI FHRDCK 660  
 Db 617 LOYFGCSIHGQDLNEDGLIDLAVGALGNVILWSPVVOINASLHFFPSKINI FHRDCK 676  
 Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGI 691  
 Db 677 RSGRDATCLAAFLCFTPIFLAPHFQTTTGI 707

RESULT 19

ADC10845

ID ADC10845 standard; protein; 707 AA.

AC ADC10845;

XX

DT 18-DEC-2003 (first entry)

DE Human extracellular matrix protein from gene 38.

XX Extracellular matrix protein; cytostatic; antibacterial; virucide;  
 KW neuroprotective; gynaecological; gastrointestinal-Gen; cardiant;  
 KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;  
 KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;  
 KW neotropic; anti-allergic; cancer; bacterial infection; viral infection;  
 KW neural disorder; immune system disorder; blood disorder;  
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
 KW pulmonary disorder; cardiovascular disorder; renal disorder;  
 KW inflammatory disorder; proliferative disorder; Human.

XX Homo sapiens.

XX US2003059875-A1.

XX 27-MAR-2003.

XX

XX

XX

PF 19-APR-2002; 2002US-00125540.

XX

PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215133P.

PR 07-JUL-2000; 2000US-0216647P.

PR 11-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.



PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-02255114P.  
PR 14-AUG-2000; 2000US-02255266P.  
PR 14-AUG-2000; 2000US-0225567P.  
PR 14-AUG-2000; 2000US-0225568P.  
PR 14-AUG-2000; 2000US-0225570P.  
PR 14-AUG-2000; 2000US-0225571P.  
PR 14-AUG-2000; 2000US-0225575P.  
PR 14-AUG-2000; 2000US-0225578P.  
PR 14-AUG-2000; 2000US-0225579P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226682P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0226924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229309P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231342P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234597P.  
PR 25-SEP-2000; 2000US-0234988P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235634P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239340P.  
PR 13-OCT-2000; 2000US-0239353P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0251990P.  
PR 05-JAN-2001; 2000US-0254097P.  
PR 17-JAN-2001; 2001US-0259678P.  
XX 17-JAN-2001; 2001US-00764870.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Ruben SM, Barash SC;  
XX WPI: 2003-743765/70.  
XX N-PSDB; ADC10580.  
PT New isolated nucleic acids and polypeptides, useful for diagnosing, treating, and/or preventing disorders, such as cancer, infections, cardiovascular and inflammatory diseases.  
PT Claim 11; SEQ ID NO 313; 235pp; English.  
XX The invention relates to an isolated nucleic acid molecule (cDNA) encoding a human extracellular matrix protein, representing one of 161 novel genes. Also included are recombinant vectors, host cells (expressing the protein), the extracellular matrix proteins (including their fragments, epitopes and homologues), an isolated antibody that binds specifically to the protein, diagnosing a pathological condition or susceptibility to a pathological condition (comprising determining the presence or absence of a mutation in the nucleic acid and diagnosing a condition based on the presence or absence of the mutation), diagnosing a pathological condition or susceptibility to a pathological condition (comprising determining the presence or amount of expression of the protein in a biological sample and diagnosing a condition based on the presence or amount of expression of the protein), preventing, treating or ameliorating a medical condition by administering the nucleic acid or protein to a mammalian subject, identifying a binding partner to the

CC protein, the gene corresponding to the cDNA sequence, and identifying an  
CC activity in a biological assay (comprising expressing the nucleic acid in  
CC a cell, isolating the supernatant, detecting an activity in a biological  
CC assay and identifying the protein in the supernatant having the  
CC activity). The nucleic acids and proteins display the following  
CC activities: Cytostatic, antibacterial, virucide, Neuroprotective,  
CC Gynaecological, Gastrointestinal-Gen, Cardiant, Cardiovascular-Gen,  
CC Nephrotropic, Antiinflammatory, Muscular-Gen, Respiratory-Gen,

Query Match 41.2%; Score 489; DB 7; Length 707;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MDLPRGLVAVWALSLPGFTDTFNMTRKPRVPGSTAPGTYVQOHDISGNKMLVVG 60  
Db 17 MDLPRGLVAVWALSLPGFTDTFNMTRKPRVPGSTAPGTYVQOHDISGNKMLVVG 76  
QY 61 PLETNGYQKTDVYKCPVIRHNCCTKLNLGRVTLNSVSRKDNMRGLSLATNPKNFLA 120  
Db 77 PLETNGYQKTDVYKCPVIRHNCCTKLNLGRVTLNSVSRKDNMRGLSLATNPKNFLA 136  
QY 121 CSPLWSHECGSSVYTCMCGRVNSRPFSTKVPALQRCQTYMDIVTLVLDGNSIYPWYE 180  
Db 137 CSPLWSHECGSSVYTCMCGRVNSRPFSTKVPALQRCQTYMDIVTLVLDGNSIYPWVE 196  
QY 181 VOHFLINILKFFYIGPGQIQGVVQYGEDVVFHFLMDYRSVKDVEAASHIEQRGGTET 240  
Db 197 VOHFLINILKFFYIGPGQIQGVVQYGEDVVFHFLMDYRSVKDVEAASHIEQRGGTET 256  
QY 241 RTAPGIEFARSEAFQKGRGAKKVMIVITDGHSHQSPDLKVIQOSERNVTRYAVVL 300  
Db 257 RTAPGIEFARSEAFQKGRGAKKVMIVITDGHSHQSPDLKVIQOSERNVTRYAVVL 316  
QY 301 GYNRRGINPETFLNEIKYIASDPDDKHFFNVITDEAALKDIVDALGDRIFSLGTTKNET 360  
Db 317 GYNRRGINPETFLNEIKYIASDPDDKHFFNVITDEAALKDIVDALGDRIFSLGTTKNET 376  
QY 361 SFGLMSQTFSSHVDGVLGAVGYDNGAVLKEKTSAGKVIPLRESYLKEFPBELKN 420  
Db 377 SFGLMSQTFSSHVDGVLGAVGYDNGAVLKEKTSAGKVIPLRESYLKEFPBELKN 436  
QY 421 HGAYLGVTVTSVSSRQGRVYVAGAPFNHTGKVIPLTMNNSLTTHQAMRQQIGSYF 480  
Db 437 HGAYLGVTVTSVSSRQGRVYVAGAPFNHTGKVIPLTMNNSLTTHQAMRQQIGSYF 496  
QY 481 GSEITSDIDGQVTDVLLVAGAPYFNEGRERKGVVYVELQNRFPVYNGTLKDSHSYQNA 540  
Db 497 GSEITSDIDGQVTDVLLVAGAPYFNEGRERKGVVYVELQNRFPVYNGTLKDSHSYQNA 556  
QY 541 RFGSSIASVRDLNODSNDVWVGAPLEDNHAGAIYIPHGFRGSILKTPKORITASELATG 600  
Db 557 RFGSSIASVRDLNODSNDVWVGAPLEDNHAGAIYIPHGFRGSILKTPKORITASELATG 616  
QY 601 LQYFGCSIHQGLDLNEDGLIDLAVGALGNVILMSRPVQINASLHFEPSKINIFHRDCK 660  
Db 617 LQYFGCSIHQGLDLNEDGLIDLAVGALGNVILMSRPVQINASLHFEPSKINIFHRDCK 676  
QY 661 RSGRDATCLAAFLCFTPIFLAPHQTTTGGI 691  
Db 677 RSGRDATCLAAFLCFTPIFLAPHQTTTGGI 707

RESULT 20  
AAU76854  
ID AAU76854 standard; protein; 193 AA.  
AC AAU76854;  
XX  
XX 21-MAY-2002 (first entry)  
XX Human integrin alpha subunit Alpha 11 A domain.  
XX Human; integrin alpha subunit; A domain; Alpha 11; integrin beta subunit;

KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;  
KW ischaemia-reperfusion injury; immune complex; parasitic disease;  
KW antiinflammatory; vasotropic; antiparasitic; vulnery; gene therapy.  
XX  
OS Homo sapiens.  
PN WO200209737-A1.  
XX  
PD 07-FEB-2002.  
XX  
XX 31-JUL-2001; 2001WO-US023957.  
XX  
PR 31-JUL-2000; 2000US-0221950P.  
PR 11-JAN-2001; 2001US-00758493.  
PR 13-MAR-2001; 2001US-00805354.  
XX  
PA (GEO) GEN HOSPITAL CORP.  
XX  
XX Arnaout AM, Li R, Xiong J;  
XX WPI; 2002-188687/24.  
XX  
XX Novel high affinity integrin polypeptide useful for treating restenosis  
PT and parasitic diseases, comprises all or part of variant integrin alpha  
PT subunit A domain or variant integrin beta subunit A-like domain.  
XX  
XX Example 2; Fig 5; 55pp; English.

CC The invention relates to a high affinity integrin polypeptide comprising  
CC all or part of a variant integrin alpha subunit A domain or a variant  
CC integrin beta subunit A-like domain. The polypeptide, preferably the  
CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by  
CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or  
CC V at residue 315 and A at residue 320 have been replaced by C, is useful  
CC for determining if a test compound is a candidate compound for binding to  
CC CD11b or for treating an inflammatory disorder, by contacting a test  
CC compound with the polypeptide and determining if the test compound binds  
CC to the polypeptide. The integrin subunits are useful for reducing  
CC skeletal muscle injury, for treating disorders caused by ischaemia-  
CC reperfusion injury, immune complexes, restenosis and parasitic diseases,  
CC to purify variant integrin polypeptide ligands and as bait proteins in  
CC two-hybrid or three-hybrid assays. This sequence represents the human  
CC integrin alpha subunit Alpha 11 A domain

QY Sequence 193 AA;  
Query Match 16.2%; Score 193; DB 5; Length 193;  
Best Local Similarity 100.0%; Pred. No. 2.1e-182;  
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 CQTYMDIVVLVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVFHFLND 218  
Db 1 CQTYMDIVVLVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVFHFLND 60  
QY 219 YRSVKDVEAASHIEQRGGTETRTAFGIEFARSEAFQKGRGAKKVMIVITDGHSHQSP 278  
Db 61 YRSVKDVEAASHIEQRGGTETRTAFGIEFARSEAFQKGRGAKKVMIVITDGHSHQSP 120  
QY 279 DLEKVIQOSERNVTRYAVAVLGYNNRRGINPETFLNEIKYIASDPDDKHFPNVTDEAL 338  
Db 121 DLEKVIQOSERNVTRYAVAVLGYNNRRGINPETFLNEIKYIASDPDDKHFPNVTDEAL 180  
QY 339 KDIVDALGDRIFS 351  
Db 181 KDIVDALGDRIFS 193  
RESULT 21  
AAU76863  
ID AAU76863 standard; protein; 193 AA.  
XX  
XX AC AAU76863;

DT 21-MAY-2002 (first entry)  
DE Human integrin alpha subunit Alpha 11 variant A domain.  
XX  
XX Human; integrin alpha subunit; A domain; Alpha 11; integrin beta subunit;  
KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;  
KW ischaemia-reperfusion injury; immune complex; parasitic disease; mutant;  
KW antiinflammatory; vasotropic; antiparasitic; vulnerary; gene therapy;  
KW mutein.  
XX  
OS Homo sapiens.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Misc-difference 191  
FT /note= "Wild-type Ile substituted by any other amino  
FT acid"  
XX  
XX WO200209737-A1.  
XX  
XX 07-FEB-2002.  
XX  
XX 31-JUL-2001; 2001WO-US023957.  
XX  
XX 31-JUL-2000; 2000US-0221950P.  
PR 11-JAN-2001; 2001US-00758493.  
PR 13-MAR-2001; 2001US-00805354.  
XX  
XX (GENO ) GEN HOSPITAL CORP.  
XX  
XX Arnaout AM, Li R, Xiong J;  
XX  
XX WPI; 2002-188687/24.  
DR  
XX Novel high affinity integrin polypeptide useful for treating restenosis  
XX and parasitic diseases, comprises all or part of variant integrin alpha  
XX subunit A domain or variant integrin beta subunit A-like domain.  
XX  
XX Claim 57; Page: 55pp; English.  
XX  
XX The invention relates to a high affinity integrin polypeptide comprising  
XX all or part of a variant integrin alpha subunit A domain or a variant  
XX integrin beta subunit A-like domain. The polypeptide, preferably the  
XX CD11b alpha subunit A domain, where I at residue 332 has been replaced by  
XX G or A, F at residue 313 and A at residue 320 have been replaced by C, or  
XX V at residue 315 and A at residue 320 have been replaced by C, is useful  
XX for determining if a test compound is a candidate compound for binding to  
XX CD11b or for treating an inflammatory disorder, by contacting a test  
XX compound with the polypeptide and determining if the test compound binds  
XX to the polypeptide. The integrin subunits are useful for reducing  
XX skeletal muscle injury, for treating disorders caused by ischaemia-  
XX reperfusion injury, immune complexes, restenosis and parasitic diseases,  
XX to purify variant integrin polypeptide ligands and as bait proteins in  
XX two-hybrid or three-hybrid assays. This sequence represents a human  
XX integrin alpha subunit Alpha 11 variant A domain. Note: This variant  
XX sequence is not featured in the specification but has been derived from  
XX the wild-type protein shown in AAU76854  
XX  
XX Sequence 193 AA;  
SQ  
Query Match 16.0%; Score 190; DB 5; Length 193;  
Best Local Similarity 100.0%; Pred. No. 2e-179;  
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 159 CQTMDIVIVLDGNSITPWEVQHFLINILKRFYIGPGQIQGVQYGVDFVHFLND 218  
DB 1 CQTMDIVIVLDGNSITPWEVQHFLINILKRFYIGPGQIQGVQYGVDFVHFLND 60  
QY 219 YRSVKDVEAASHIEQGGTETRTAFGIEFARSEAFQGGKAKKVMIVITDGHSDSP 278  
DB 61 YRSVKDVEAASHIEQGGTETRTAFGIEFARSEAFQGGKAKKVMIVITDGHSDSP 120  
QY 279 DLEKVIQSSRDNVTRYAVAVLGYNNRRGINPETFLNEIKYIASDDPKHFFNVTDEAL 338

DB 121 DLEKVIQSSRDNVTRYAVAVLGYNNRRGINPETFLNEIKYIASDDPKHFFNVTDEAL 180  
QY 339 KDIVDALGDR 348  
DB 181 KDIVDALGDR 190  
RESULT 22  
AAB50087  
ID AAB50087 standard; protein; 1188 AA.  
XX  
XX AAB50087;  
AC  
XX  
XX 19-MAR-2001 (first entry)  
DT  
XX  
DE Murine A259.  
XX  
XX Human; A259; integrin alpha subunit; fibrosis; liver disease; lung;  
KW kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;  
KW rheumatoid arthritis.  
XX  
XX Mus sp.  
XX  
XX Key Location/Qualifiers  
FH Domain 1..1141  
FT /label= Extracellular\_domain  
FT Peptide 1..22  
FT /label= Signal\_peptide  
FT Protein 23..1188  
FT /label= Mature\_protein  
FT Domain 39..74  
FT /label= Integrin\_alphasubunit\_repeat\_domain\_#1  
FT Domain 115..157  
FT /label= Integrin\_alphasubunit\_repeat\_domain\_#2  
FT Domain 164..345  
FT /label= I\_domain  
FT Domain 367..392  
FT /label= Integrin\_alphasubunit\_repeat\_domain\_#3  
FT Domain 421..455  
FT /label= Integrin\_alphasubunit\_repeat\_domain\_#4  
FT Domain 478..516  
FT /label= Integrin\_alphasubunit\_repeat\_domain\_#5  
FT Domain 540..575  
FT /label= Integrin\_alphasubunit\_repeat\_domain\_#6  
FT Domain 602..640  
FT /label= Integrin\_alphasubunit\_repeat\_domain\_#7  
FT Domain 1142..1164  
FT /label= Transmembrane\_domain  
FT Domain 1165..1188  
FT /label= Cytoplasmic\_domain  
XX  
XX WO200073339-A1.  
XX  
XX 07-DEC-2000.  
XX  
XX 15-MAY-2000; 2000WO-US013262.  
XX  
XX 28-MAY-1999; 99US-00322790.  
XX 27-APR-2000; 2000US-00561263.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Pan Y, Lora JM;  
XX  
XX WPI; 2001-041142/05.  
DR N-PSDB; AAC91904, AAC91905.  
XX  
XX Nucleic acid encoding alpha-integrin subunits, useful for treatment and  
XX diagnosis of fibrosis, e.g. of the liver.  
XX  
XX Claim 8; Fig 5; 164pp; English.  
XX





XX AAU19822;  
AC 06-DEC-2001 (first entry)  
DT Human novel extracellular matrix protein, Seq ID No 472.  
DE  
XX  
DE  
XX  
XX Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;  
KW antianemic; antirheumatic; antisclerotic; cardiant; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;  
KW antialzheimers; immune/autimmune disease; HIV infection; anaemia;  
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;  
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;  
KW Sezary syndrome; Gaucher's disease; neurological diseases;  
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
KW cardiac arrest; tachycardia; angina; infection; corneal infections;  
KW wound healing; immunogen; gene therapy; antisense; food additive.  
XX  
OS Homo sapiens.  
XX  
XX WO200155368-A1.  
PN  
XX  
XX 02-AUG-2001.  
PD  
XX  
XX 17-JAN-2001; 2001WO-US001348.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205115P.  
PR 01-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0232968P.  
PR 14-SEP-2000; 2000US-0232977P.  
PR 14-SEP-2000; 2000US-0232988P.  
PR 14-SEP-2000; 2000US-0232999P.  
PR 14-SEP-2000; 2000US-0233400P.  
PR 14-SEP-2000; 2000US-0233401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.

PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251398P.  
 PR 06-DEC-2000; 2000US-0256719P.  
 PR 08-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 11-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX N-PSDB; AAS31393.  
 DR WPI; 2001-465572/50.  
 DR N-PSDB; AAS31393.  
 XX  
 PT Nucleic acid molecules encoding human secreted extracellular matrix  
 PT proteins, used in preventing, treating or ameliorating a disorder, e.g.  
 PT Alzheimer's and Parkinson's diseases and cancers.  
 XX  
 PS Claim 11; SEQ ID NO 472; 577pp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding novel  
 CC human secreted extracellular matrix proteins (SPs). The polynucleotides  
 CC and proteins are used to prevent, treat a medical condition in e.g.  
 CC humans, mice, rabbits, goats, cats, dogs, chickens or sheep. For  
 CC example, disorders associated with decreased expression of SPs. The SP  
 CC polynucleotide or a vector expressing them may be administered to treat  
 CC diseases by gene therapy. Antisense molecules may be administered to down  
 CC regulate expression of SPs by binding with the cells own genes and  
 CC preventing their expression. The polynucleotides may also be used as DNA  
 CC probes in diagnostic assays. The SPs may also be used as antigens to  
 CC produce antibodies and to identify modulators (agonists and antagonists)  
 CC of the SPs. The anti-(SP) antibodies and antagonists may also be used to  
 CC down regulate expression and activity of SP and as diagnostic agents for  
 CC detecting the presence of SPs in samples. The disorders include for  
 CC example: immun/autoimmune diseases (e.g. HIV (human immunodeficiency  
 CC virus) infections, anaemia, rheumatoid arthritis and multiple sclerosis),  
 CC cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of  
 CC the breast or liver, Sezary syndrome and Gaucher's disease), neurological  
 CC diseases (e.g. Alzheimer's disease, Parkinson's disease), cardio-  
 CC /cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina),  
 CC infections caused by bacteria, viruses and fungi and ocular disorders  
 CC (e.g. corneal infections). Other uses include wound healing, maintenance  
 CC of organs before transplantation, support of cell culture of primary

Query Match 7.4%; Score 88; DB 4; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-78;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 CQYMDIVIVLDCGNSIYFWVQVHFLNLIKFKYIGPGQIQGVVQVCGEDVYHREHND 218  
 |||||||  
 Db 8 CQYMDIVIVLDCGNSIYFWVQVHFLNLIKFKYIGPGQIQGVVQVCGEDVYHREHND 67  
 |||||||

QY 219 YRSVKDWEAAASHIEQGGTETRTAFGI 246  
 |||||||  
 Db 68 YRSVKDWEAAASHIEQGGTETRTAFGI 95  
 |||||||

RESULT 27  
 AAU87675  
 ID AAU87675 standard; protein; 103 AA.  
 XX  
 AC AAU87675;  
 XX

DT 05-JUN-2002 (first entry)  
 XX  
 DE Novel central nervous system protein #585.  
 XX  
 KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
 KW hyperproliferative disorder; neoplasm; cardiovascular disorder;  
 KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;  
 KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
 KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
 KW adenocarcinoma; reproductive system disorder; testicular feminisation;  
 KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;  
 KW respiratory disorder; renal disorder; kidney failure; blood disorder;  
 KW myocardial infarction; wound healing; cell proliferation; skin aging;  
 KW food additive; food preservative; gene therapy.  
 OS Homo sapiens.  
 XX  
 XX WO200155318-A2.  
 PN  
 XX  
 PD 02-AUG-2001.  
 XX  
 XX  
 PF 17-JAN-2001; 2001WO-US001332.  
 XX  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214888P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226682P.  
 PR 23-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.

PR	08-SEP-2000	2000US-02320811
PR	12-SEP-2000	2000US-02321968P
PR	14-SEP-2000	2000US-02323977P
PR	14-SEP-2000	2000US-02323988P
PR	14-SEP-2000	2000US-02323998P
PR	14-SEP-2000	2000US-02324007P
PR	14-SEP-2000	2000US-02324019P
PR	14-SEP-2000	2000US-02324033P
PR	14-SEP-2000	2000US-02330649P
PR	14-SEP-2000	2000US-02330659P
PR	21-SEP-2000	2000US-02342234P
PR	21-SEP-2000	2000US-02342243P
PR	25-SEP-2000	2000US-02345979P
PR	25-SEP-2000	2000US-02345988P
PR	25-SEP-2000	2000US-02354849P
PR	27-SEP-2000	2000US-02358349P
PR	27-SEP-2000	2000US-02358369P
PR	29-SEP-2000	2000US-02363277P
PR	29-SEP-2000	2000US-02363278P
PR	29-SEP-2000	2000US-02363689P
PR	29-SEP-2000	2000US-02363698P
PR	29-SEP-2000	2000US-02363709P
PR	13-OCT-2000	2000US-02370409P
PR	13-OCT-2000	2000US-02370409P
PR	13-OCT-2000	2000US-02395359P
PR	20-OCT-2000	2000US-02399379P
PR	20-OCT-2000	2000US-02418089P
PR	20-OCT-2000	2000US-02418099P
PR	20-OCT-2000	2000US-02418269P
PR	20-OCT-2000	2000US-02446179P
PR	08-NOV-2000	2000US-02464749P
PR	08-NOV-2000	2000US-02464759P
PR	08-NOV-2000	2000US-02464769P
PR	08-NOV-2000	2000US-02464779P
PR	08-NOV-2000	2000US-02464789P
PR	08-NOV-2000	2000US-02465239P
PR	08-NOV-2000	2000US-02465249P
PR	08-NOV-2000	2000US-02465259P
PR	08-NOV-2000	2000US-02465269P
PR	08-NOV-2000	2000US-02465279P
PR	08-NOV-2000	2000US-02465289P
PR	08-NOV-2000	2000US-02465329P
PR	08-NOV-2000	2000US-02466099P
PR	08-NOV-2000	2000US-02466109P
PR	08-NOV-2000	2000US-02466119P
PR	17-NOV-2000	2000US-02492129P
PR	17-NOV-2000	2000US-02492139P
PR	17-NOV-2000	2000US-02492149P
PR	17-NOV-2000	2000US-02492159P
PR	17-NOV-2000	2000US-02492169P
PR	17-NOV-2000	2000US-02492179P
PR	17-NOV-2000	2000US-02492189P
PR	17-NOV-2000	2000US-02492449P
PR	17-NOV-2000	2000US-02492459P
PR	17-NOV-2000	2000US-02492649P
PR	17-NOV-2000	2000US-02492659P
PR	17-NOV-2000	2000US-02492979P
PR	17-NOV-2000	2000US-02492999P
PR	17-NOV-2000	2000US-02493009P

01-DEC-2000; 2000US-0250160P.  
01-DEC-2000; 2000US-0250391P.  
05-DEC-2000; 2000US-0251030P.  
05-DEC-2000; 2000US-0251988P.  
05-DEC-2000; 2000US-0256719P.  
06-DEC-2000; 2000US-0251479P.  
08-DEC-2000; 2000US-0251856P.  
08-DEC-2000; 2000US-0251869P.  
08-DEC-2000; 2000US-0251989P.  
08-DEC-2000; 2000US-0251990P.  
11-DEC-2000; 2000US-0254097P.  
05-JAN-2001; 2001US-0259678P.  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-581633/65.  
N-PSDB; ABK44005.  
New isolated nucleic acid encoding a protein for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.  
Claim 9; SEQ ID NO 1193; 837pp; English.  
The invention describes an isolated nucleic acid molecule (I) encoding a novel central nervous system protein. (i) and polypeptides (iii) encoded by (I), are used to treat a medical conditions and in diagnosis of a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease and amyotrophic lateral sclerosis, infections caused by bacteria, viruses e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders e.g. corneal infection, gastrointestinal disorders e.g. dysphagia, adenocarcinomas and irritable bowel syndrome, reproductive system disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes and pituitary dwarfism, cancers and disorders at the cellular level e.g. leukaemia, disorders involving neovascularisation e.g. malignancies, respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g. acute kidney failure and blood related disorders e.g. myocardial infarction. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein,

Query Match	7.4%	Score 88	DB 4	Length 103
Best Local Similarity	100.0%	Pred. No. 1.9e-78		
Matches 88	Conservative	0	Mismatches 0	Indels 0
Gaps				0

Qy	159	CQTMDIVIVL	DGNSNSIY	PMVVEQHF	INILK	KPYIGFG	QIQVGW	QYQGEDV	WHF	HLND	218
Dd	8	CQTMDIVIVL	DGNSNSIY	PMVVEQHF	INILK	KPYIGFG	QIQVGW	QYQGEDV	WHF	HLND	67
Qy	219	YRSKDV	WEAASHIE	QRGGTET	TFACI	246					
Dd	68	YRSKDV	WEAASHIE	QRGGTET	TFACI	95					

```

RESULT 28
ABP48042
ID ABP48042 standard; protein; 103 AA.
XX
XX
XX AC
XX AC
XX
XX
XX 23-AUG-2002 (first entry)
XX
XX
DE Human polypeptide SEQ ID NO 472.

```





PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234224P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241825P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 17-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0255719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251855P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2000US-0259678P.  
PR 17-JAN-2001; 2000US-00764870.

{HUMA-} HUMAN GENOME SCI INC.  
Rosen CA, Ruben SM, Barash SC;  
WPI; 2003-743765/70.  
N-PSDB; ADC10739.

New isolated nucleic acids and polypeptides, useful for diagnosing, treating, and/or preventing disorders, such as cancer, infections, cardiovascular and inflammatory diseases.

XX Claim 11; SEQ ID NO 472; 235pp; English.

PS The invention relates to an isolated nucleic acid molecule (cDNA)

XX encoding a human extracellular matrix protein, representing one of 161

CC novel genes. Also included are recombinant vectors, host cells

CC (expressing the protein), the extracellular matrix proteins (including

CC their fragments, epitopes and homologues), an isolated antibody that

CC binds specifically to the protein, diagnosing a pathological condition or

CC susceptibility to a pathological condition (comprising determining the

CC presence or absence of a mutation in the nucleic acid and diagnosing a

CC condition based on the presence or absence of the mutation), diagnosing a

CC pathological condition or susceptibility to a pathological condition

CC (comprising determining the presence or amount of expression of the

CC protein in a biological sample and diagnosing a condition based on the

CC presence or amount of expression of the protein), preventing, treating or

CC ameliorating a medical condition by administering the nucleic acid or

CC protein to a mammalian subject, identifying a binding partner to the

CC protein, the gene corresponding to the cDNA sequence, and identifying an

CC activity in a biological assay (comprising expressing the nucleic acid in

CC a cell, isolating the supernatant, detecting an activity in a biological

CC assay and identifying the protein in the supernatant having the

CC activity). The nucleic acids and proteins display the following

CC activities: Cytostatic, antibacterial, Virucide, Neuroprotective,

CC Gynaecological, Gastrointestinal-Gen, Cardiant, Cardiovascular-Gen,

CC Nephrotropic, Antiinflammatory, Muscular-Gen, Respiratory-Gen,

Query Match 7.4%; Score 88; DB 7; Length 103;

Best Local Similarity 100.0%; Pred. No. 1.9e-78;

Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 CQTYMDIVLVGSGNSIYVWVEVQHFLINILKFFYIGPQIQVGVVQGEDVYVHFLND 218

DB 8 CQTYMDIVLVGSGNSIYVWVEVQHFLINILKFFYIGPQIQVGVVQGEDVYVHFLND 67

QY 219 YRSVKDVAASHIEORGCTGTARGI 246

DB 68 YRSVKDVAASHIEORGCTGTARGI 95

RESULT 30

ADB32059

ID ADB32059 standard; peptide; 42 AA.

AC ADB32059;

XX 04-DEC-2003 (first entry)

DE alphaA-integrin alpha subunit alpha1.

XX Integrin; alphaV-beta3 integrin; RGD peptide ligand; modulator; agonist;

KW antagonist; alpha1.

XX Unidentified.

OS WO2003067219-A2.

PN 14-AUG-2003.

XX 07-FEB-2003; 2003WO-US003903.

PF 07-FEB-2002; 2002US-0354773P.

XX (GEO ) GEN HOSPITAL CORP.

PA Arnaout AM;

XX WPI; 2003-663639/G2.

DR Screening potential modulators of alphaVbeta3 integrin useful to identify

XX agonists and antagonists uses computer model of three-dimensional

PT structure including a binding site and data from an alphaVbeta3 integrin-

PT ligand complex.

XX Disclosure; Page 5c; 49pp; English.

PS The invention relates to a method for screening test compounds as

XX potential modulators of alphaV-beta3 integrin using a computer model of

CC the three-dimensional structure of alphaV-beta3 integrin which includes a

CC binding site. The model is based on atomic coordinates of defined alphaV-

CC beta3 integrin amino acids obtained from the structure of a complex of

CC alphaV-beta3 integrin with a known 'RGD peptide' ligand. The method is

CC useful to identify alphaV-beta3 integrin ligands that, because they bind

CC to alphaV-beta3 integrin, may be modulators e.g. agonists or antagonist

CC of alphaV-beta3 integrin activity. It is useful to ascertain whether a

CC specific test compound is a potential modulator and especially to greatly

CC reduce numbers of compounds which must be further tested for their

CC ability to modulate alphaV-beta3 integrin activity. The current sequence

CC represents the alphaA-integrin alpha subunit alpha1.

XX Sequence 42 AA;

QY 335 EAALKDIDVALGDRIFSLGNTKNETSGLEMSQTGSSHHV 376

DB 1 EAALKDIDVALGDRIFSLGNTKNETSGLEMSQTGSSHHV 42

RESULT 31

ABB72288

ID ABB72288 standard; protein; 545 AA.

XX AC ABB72288;

XX 04-APR-2002 (first entry)

DE Murine protein isolated from skin cells SEQ ID NO: 500.

XX Human; rat; mouse; skin cell; skin wound; cancer; growth defect;

KW developmental defect; inflammatory disease; dermatological; vulnerary;

KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

XX Mus sp.

OS WO200190357-A1.

FN 29-NOV-2001.

XX 24-MAY-2001; 2001WO-NZ000099.

PF 24-MAY-2000; 2000US-0206650P.

PR 25-JUL-2000; 2000US-0221232P.

XX (GENE-) GENESIS RBS & DEV CORP LTD.

PA Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;

PI Kumble KD;

XX WPI; 2002-122020/16.

DR New polynucleotides and polypeptides encoded by the polynucleotides

XX isolated from skin cells, useful for treating skin wounds, cancers,

PT growth and developmental defects, inflammatory diseases, or for

PT modulating immune responses.

PS Claim 4; Page 305-306; 466pp; English.

XX The present invention provides the protein and coding sequences of cDNAs

CC isolated from human, murine and rat skin cell libraries. The sequences

CC can be used in the development of therapeutic agents useful in the

CC treatment of skin diseases, including skin wounds, cancer, growth

CC defects, developmental defects and inflammatory diseases. The proteins

CC have important roles in the induction of hair growth, cell proliferation

CC and cell-cell interaction, in maintaining tissue integrity, in wound  
CC healing and in modulating immune responses. The present sequence is a  
CC polypeptide of the invention

SQ Sequence 545 AA;

Query Match 3.5%; Score 42; DB 5; Length 545;  
Best Local Similarity 100.0%; Pred. No. 3.5e-32;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYCO 807  
123 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYCO 164

RESULT 32

ABB72300  
ID ABB72300 standard; protein; 688 AA.

XX AC ABB72300;

XX DT 04-APR-2002 (first entry)

XX DE Rat protein isolated from skin cells SEQ ID NO: 624.

XX KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;  
XX KW developmental defect; inflammatory disease; dermatological; vulnary;  
XX KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

XX OS Rattus sp.

XX FN WO200190357-A1.

XX PD 29-NOV-2001.

XX PF 24-MAY-2001; 2001WO-NZ000099.

XX PR 24-MAY-2000; 2000US-0206650P.

XX PR 25-JUL-2000; 2000US-0221232P.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PI Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;

XX PI Kumble KD;

XX DR WPI; 2002-122020/16.

XX DR N-ESDB; ABL34985.

XX PT New polynucleotides and polypeptides encoded by the polynucleotides  
XX PT isolated from skin cells, useful for treating skin wounds, cancers,  
XX PT growth and developmental defects, inflammatory diseases, or for  
XX PT modulating immune responses.

XX FS Claim 4; Page 388-390; 466pp; English.

XX CC The present invention provides the protein and coding sequences of cDNAs  
XX CC isolated from human, murine and rat skin cell libraries. The sequences  
XX CC can be used in the development of therapeutic agents useful in the  
XX CC treatment of skin diseases, including skin wounds, cancer, growth  
XX CC defects, developmental defects and inflammatory diseases. The proteins  
XX CC have important roles in the induction of hair growth, cell proliferation  
XX CC and cell-cell interaction, in maintaining tissue integrity, in wound  
XX CC healing and in modulating immune responses. The present sequence is a  
XX CC polypeptide of the invention

XX SQ Sequence 688 AA;

Query Match 3.5%; Score 42; DB 5; Length 688;  
Best Local Similarity 100.0%; Pred. No. 4.3e-32;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYCO 807  
|||||

Db 266 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYCO 307

RESULT 33

ABB72289

ID ABB72289 standard; protein; 696 AA.

XX AC ABB72289;

XX DT 04-APR-2002 (first entry)

XX DE Rat protein isolated from skin cells SEQ ID NO: 501.

XX KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;  
XX KW developmental defect; inflammatory disease; dermatological; vulnary;  
XX KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

XX OS Rattus sp.

XX PN WO200190357-A1.

XX PD 29-NOV-2001.

XX PF 24-MAY-2001; 2001WO-NZ000099.

XX PR 24-MAY-2000; 2000US-0206650P.

XX PR 25-JUL-2000; 2000US-0221232P.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PI Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;

XX PI Kumble KD;

XX DR WPI; 2002-122020/16.

XX PT New polynucleotides and polypeptides encoded by the polynucleotides  
XX PT isolated from skin cells, useful for treating skin wounds, cancers,  
XX PT growth and developmental defects, inflammatory diseases, or for  
XX PT modulating immune responses.

XX PS Claim 4; Page 306-307; 466pp; English.

XX CC The present invention provides the protein and coding sequences of cDNAs  
XX CC isolated from human, murine and rat skin cell libraries. The sequences  
XX CC can be used in the development of therapeutic agents useful in the  
XX CC treatment of skin diseases, including skin wounds, cancer, growth  
XX CC defects, developmental defects and inflammatory diseases. The proteins  
XX CC have important roles in the induction of hair growth, cell proliferation  
XX CC and cell-cell interaction, in maintaining tissue integrity, in wound  
XX CC healing and in modulating immune responses. The present sequence is a  
XX CC polypeptide of the invention

XX SQ Sequence 696 AA;

Query Match 3.5%; Score 42; DB 5; Length 696;  
Best Local Similarity 100.0%; Pred. No. 4.3e-32;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYCO 807

Db 266 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYCO 307  
|||||

RESULT 34

AA39928

ID AA39928 standard; protein; 757 AA.

XX AC AA39928;

XX DT 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 3073.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
XX WO200153312-A1.  
XX  
XX  
XX PD 26-JUL-2001.  
XX  
XX PF 26-DEC-2000; 2000NO-US034263.  
XX  
XX PR 23-DEC-1999; 99US-00471275.  
XX PR 21-JAN-2000; 2000US-00488725.  
XX PR 25-APR-2000; 2000US-00552317.  
XX PR 20-JUN-2000; 2000US-00598042.  
XX PR 19-JUL-2000; 2000US-00620312.  
XX PR 03-AUG-2000; 2000US-00653450.  
XX PR 14-SEP-2000; 2000US-00662191.  
XX PR 19-OCT-2000; 2000US-00693036.  
XX PR 29-NOV-2000; 2000US-00727344.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
XX PI Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI; 2001-442253/47.  
XX N-PSDB; AA159084.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders such  
XX as central nervous system injuries.  
XX  
XX Example 4; SEQ ID NO 3073; 10078pp; English.  
XX  
XX The invention relates to human nucleic acids (AA157798-AA161369) and the  
XX encoded polypeptides (AA33642-AA42213) with nootropic, and the  
XX immunosuppressant and cytostatic activity. The polynucleotides are useful  
XX in gene therapy. A composition containing a polypeptide or polynucleotide  
XX of the invention may be used to treat diseases of the peripheral nervous  
XX system, such as peripheral nervous injuries, peripheral neuropathy and  
XX localised neuropathies and central nervous system diseases, such as  
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
XX utilisation of the activities such as: immune system suppression,  
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
XX assays for receptor activity, arthritis and inflammation, leukaemias and  
XX C.N.S disorders. Note: The sequence data for this patent did not form  
XX part of the printed specification  
XX  
XX SQ Sequence 757 AA;  
Query Match 3.2%; Score 38; DB 4; Length 757;  
Best Local Similarity 100.0%; Pred. No. 4.4e-28;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 738 DTADYVKPVTFSVSEVSELDPDHGPMLDGPWTTLRVSV 775  
|||||  
DB 76 DTADYVKPVTFSVSEVSELDPDHGPMLDGPWTTLRVSV 113  
|||||  
RESULT 35  
ABG12950  
ID ABG12950 standard; protein; 117 AA.  
XX  
XX AC ABG12950;  
XX  
XX DT 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #12941.  
DE  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
KW  
XX  
OS Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
XX PD 11-OCT-2001.  
XX  
XX PF 30-MAR-2001; 2001WO-US0008631.  
XX  
XX PR 31-MAR-2000; 2000US-00540217.  
XX PR 23-AUG-2000; 2000US-00649167.  
XX (HYSE-) HYSEQ INC.  
XX  
XX PI Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
XX N-PSDB; AA577137.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 20; SEQ ID NO 43309; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
XX and in recombinant production of (II). The polynucleotides are also used  
XX in diagnostics as expressed sequence tags for identifying expressed  
XX genes. (I) is useful in gene therapy techniques to restore normal  
XX activity of (II) or to treat disease states involving (II). (II) is  
XX useful for generating antibodies against it, detecting or quantitating a  
XX polypeptide in tissue, as molecular weight markers and as a food  
XX supplement. (II) and its binding partners are useful in medical imaging  
XX of sites expressing (II). (I) and (II) are useful for treating disorders  
XX involving aberrant protein expression or biological activity. The  
XX polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
XX amino acid sequences of the invention. Note: The sequence data for this  
XX patent did not appear in the printed specification, but was obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX SQ Sequence 117 AA;  
Query Match 2.9%; Score 34; DB 4; Length 117;  
Best Local Similarity 100.0%; Pred. No. 7.9e-25;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1130 FBTSKQEDMQVPTWIIIVGSTLGLLIIALLVIAL 1163  
|||||  
DB 42 FBTSKQEDMQVPTWIIIVGSTLGLLIIALLVIAL 75  
|||||  
RESULT 36  
AAB25614  
ID AAB25614 standard; peptide; 33 AA.  
XX  
XX AC AAB25614;  
XX  
XX DT 21-NOV-2000 (first entry)  
XX  
XX DE Human secreted protein ITGA11 peptide encoded by secreted protein #6.

XX Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;  
 KW antirheumatic, dermatological; antiproliferative; antiarteriosclerotic;  
 KW anticancer; vulnary; antiviral; antibacterial; antifungal;  
 KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;  
 KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;  
 KW Crohn's disease; nephritis; hyperproliferative disorder;  
 KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;  
 KW melanoma; lymphoma; wound healing; human.  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO200029435-A1.  
 PN  
 XX  
 XX 25-MAY-2000.  
 PD  
 XX  
 XX 27-OCT-1999; 99WO-US025031.  
 PF  
 XX  
 XX 28-OCT-1998; 98US-0105971P.  
 PR  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX  
 XX Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;  
 PI Greene JM;  
 PI  
 XX WPI; 2000-387742/33.  
 DR  
 XX  
 XX Isolated nucleic acid molecules encoding human secreted proteins are used  
 XX for the prevention, amelioration and treatment of autoimmune,  
 XX inflammatory, hyperproliferative and cardiovascular disorders, cancer,  
 XX wounds, and infectious diseases.  
 XX  
 XX Disclosure; Page 107; 803pp; English.  
 PS  
 XX  
 XX The present invention relates to 12 secreted human proteins and the  
 CC nucleotide sequences encoding them. The polynucleotide sequences given in  
 CC AAB0606-A80623 encode the 12 secreted protein sequences given in  
 CC AAB25576-B25593. The human secreted proteins have various activities  
 CC dependent on the tissues in which they are expressed. Examples of the  
 CC activities of the proteins include: immunosuppressant; anti-inflammatory;  
 CC antiarthritic; antirheumatic, dermatological; antiproliferative;  
 CC antiarteriosclerotic; anticancer; vulnary; antiviral; antibacterial;  
 CC and antifungal activity. The proteins, polypeptides, agonists and  
 CC antagonists may be used to treat prevent and/or diagnose various disease,  
 CC disorders and conditions examples of which include: immune disorders e.g.  
 CC Addison's disease, rheumatoid arthritis, dermatitis, and multiple  
 CC sclerosis; inflammatory disorders e.g. inflammatory bowel disease,  
 CC Crohn's disease and nephritis; hyperproliferative disorders such as  
 CC paraproteinemias and purpura; cardiovascular disorders e.g. coronary  
 CC arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The  
 CC proteins and polynucleotide sequences may also be used in wound healing  
 CC and the treatment of infectious diseases. The human secreted protein gene  
 CC #6 and protein sequences are represented in sequences AAB0611 and  
 CC AAB2581. Sequences AAB0650-A80651 represent genes related to the  
 CC secreted protein genes#6  
 XX  
 XX Sequence 33 AA;  
 SQ  
 Query Match 2.8%; Score 33; DB 3; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-24;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 64 TNGYQKTDGVYKCPVHGNCTKLNGLGRVTLNV 96  
 Db 1 TNGYQKTDGVYKCPVHGNCTKLNGLGRVTLNV 33  
 RESULT 37  
 ADA27114  
 ID ADA27114 standard; peptide; 33 AA.  
 XX  
 AC ADA27114;  
 XX

DT 20-NOV-2003 (first entry)  
 XX Human novel secreted protein from gene 7 peptide.  
 DE  
 XX  
 XX Cytostatic; antiinflammatory; immunomodulator; neuroprotective;  
 KW hemostatic; gene therapy; cancer; inflammation; immune disorder;  
 KW neurological disorder; blood clotting disorder; food additive;  
 KW preservative; human; secreted protein.  
 XX  
 OS  
 XX Homo sapiens.  
 XX  
 XX US2003055231-A1.  
 PN  
 XX  
 XX 20-MAR-2003.  
 PD  
 XX  
 XX 29-OCT-2001; 2001US-00984130.  
 PF  
 XX  
 XX 28-OCT-1998; 98US-0105971P.  
 PR  
 XX 27-OCT-1999; 99WO-US025031.  
 PR 19-APR-2000; 2000US-0198407P.  
 PR 30-OCT-2000; 2000US-0243792P.  
 PR 18-APR-2001; 2001US-00836353.  
 XX  
 XX (NIJ/) NI J.  
 PA (YOUNG/) YOUNG P E.  
 PA (KENN/) KENNY J J.  
 PA (OLSE/) OLSEN H S.  
 PA (MOOR/) MOORE P A.  
 PA (WEI/) WEI Y.  
 PA (GREE/) GREENE J M.  
 PA (RUBE/) RUBEN S M.  
 PA (LIUD/) LIU D.  
 PA (CROC/) CROCKER P R.  
 XX  
 XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;  
 PI Ruben SM, Liu D, Crocker PR;  
 PI  
 XX WPI; 2003-567103/53.  
 DR  
 XX New human secreted nucleic acid molecules and polypeptides, useful for  
 PT preventing, treating, or ameliorating a medical condition, such as  
 PT cancer, inflammation, immune disorders, neurological and blood clotting  
 PT disorders.  
 XX  
 PS Disclosure; Page 343; 454pp; English.  
 CC  
 CC The invention relates to an isolated nucleic molecule that is at least  
 CC 95% identical to 18 human cDNA sequences representing 12 novel genes  
 CC encoding secreted proteins or a polynucleotide fragment of the cDNA  
 CC sequence contained in American Type Culture Collection (ATCC) deposit No.  
 CC defined in the specification, its species homologue, a variant or allelic  
 CC variant of the polynucleotide having a polynucleotide capable of  
 CC hybridising under conditions the polynucleotide, where the polynucleotide  
 CC does not hybridise under stringent conditions to a nucleic acid molecule  
 CC having a nucleotide sequence of only A or T residues. Also included are  
 CC recombinant vectors, host cells (for producing the polypeptide), the  
 CC secreted polypeptide (comprising a sequence that is at least 95%  
 CC identical to a polypeptide fragment, domain, epitope, full-length  
 CC protein, variant, allelic variant or species homologue), antibodies that  
 CC specifically bind to the polypeptides, diagnosing, treating, preventing  
 CC or ameliorating a medical condition by administering the polynucleotide  
 CC or the polypeptide, the gene corresponding to the cDNA sequence and  
 CC identifying an activity in a biological assay (by expressing the cDNA  
 CC sequence in a cell, isolating the supernatant, and detecting an activity  
 CC in a biological assay and identifying the protein in the supernatant  
 CC having the activity). The polypeptides, nucleic acids and antibodies are  
 CC useful for diagnosing a pathological condition or a susceptibility to a  
 CC pathological condition, for preventing, treating, or ameliorating a  
 CC medical condition, such as cancer, inflammation and other immune  
 CC disorders, neurological and blood clotting disorders (many examples are  
 CC given in the specification). The nucleic acids are also useful for  
 CC chromosome identification, radiation hybrid mapping or long-range  
 CC restriction mapping. The polypeptides and antibodies are useful for

CC providing immunological probes for differential identification of the  
 CC tissues immunohistochemistry assays. The polypeptide, polynucleotide,  
 CC agonist or antagonist may also be used as a food additive or preservative  
 CC to increase or decrease storage capabilities, fat content or other  
 CC nutritional components. The present is a peptide motif from a secreted  
 CC protein of the invention.  
 XX  
 XX Sequence 33 AA;

Query Match

Best Local Similarity 2.8%; Score 33; DB 6; Length 33;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 TNGYQKTGDYKCPVHGNCTKLNGLGRVTLNV 96  
 DB 1 TNGYQKTGDYKCPVHGNCTKLNGLGRVTLNV 33

RESULT 38

AD86651

ID AD86651 standard; peptide; 33 AA.

AC AD86651;

DT 29-JAN-2004 (first entry)

XX Transmembrane domain.

DE human; secreted protein; cancer; liver disorder; hepatitis;

KW neural disorder; Alzheimer's disease; transmembrane domain.

OS Homo sapiens.

XX US2003129685-A1.

PN 10-JUL-2003.

XX 18-APR-2001; 2001US-00836353.

XX 28-OCT-1998; 98US-0105971P.

PR 27-OCT-1999; 99WO-US025031.

PR 19-APR-2000; 2000US-0198407P.

XX (NIJ/J) NI J.

PA (YOUNG) YOUNG P E.

PA (Kenny) KENNY J J.

PA (OLSEN) OLSEN H S.

PA (MOORE) MOORE P A.

PA (WEI) WEI Y.

PA (GREENE) GREENE J M.

PA (RUBEN) RUBEN S M.

XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;

PI Ruben SM;

XX WPI; 2004-020335/02.

XX New nucleic acid molecule, useful for preparing a medicament for

PT preventing, treating or ameliorating a medical condition e.g. cancer,

PT liver disorders or neural disorders.

XX Disclosure; SEQ ID NO 102; 380pp; English.

XX The invention relates to an isolated nucleic acid sequence, or its

CC allelic variant, a fragment of the cDNA sequence, or its fragment,

CC domain, epitope or species homologue. The nucleic acid is useful for

CC preparing a medicament for preventing, treating or ameliorating a medical

CC condition e.g., cancer, liver disorders such as hepatitis or neural

CC disorders such as Alzheimer's disease. The present sequence represents

CC the amino acid sequence of a transmembrane domain.

XX Sequence 33 AA;

Query Match

Best Local Similarity 2.0%; Score 24; DB 4; Length 24;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match

Best Local Similarity 2.8%; Score 33; DB 8; Length 33;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 TNGYQKTGDYKCPVHGNCTKLNGLGRVTLNV 96  
 DB 1 TNGYQKTGDYKCPVHGNCTKLNGLGRVTLNV 33

RESULT 39

AAB30927

ID AAB30927 standard; peptide; 24 AA.

AC AAB30927;

DT 02-APR-2001 (first entry)

XX Peptide derived from a human alpha1 integrin chain.

XX Human; integrin; alpha1 subunit; fibroblast; muscle cell; chondrocyte;

KW osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;

KW wound healing; trauma; rheumatoid arthritis; osteoarthritis;

KW osteoporosis; cartilage damage; bone damage; cartilage.

OS Homo sapiens.

XX WO2000075187-A1.

XX 14-DEC-2000.

XX 31-MAY-2000; 2000WO-SE001135.

XX 03-JUN-1999; 99SE-00002056.

XX (ACTI-) ACTIVE BIOTECH AB.

XX Gullberg D;

XX WPI; 2001-071061/08.

XX Integrin subunit alpha 11 or integrin heterodimer comprising subunit

PT alpha 11 in association with subunit beta, useful for treating muscle

PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.

XX Claim 23; Page 40; 79pp; English.

XX The present sequence is derived from the cytoplasmic domain of the human

CC integrin subunit, designated alpha11. The alpha11 polynucleotide and

CC polypeptide are useful as markers of cell target molecules, such as

CC fibroblasts, muscle cells, chondrocytes, osteoblasts, mesenchymally

CC derived cells or stem cells. They are also used for determining the

CC differential-stage of cells during differentiation, development in

CC pathological conditions, in tissue regeneration, in transplantation or in

CC therapeutic and physiological repair of tissues. The pathological

CC conditions involving subunit alpha11 are selected from damage of cells,

CC muscle dystrophy, fibrosis, wound healing, trauma, rheumatoid arthritis,

CC osteoarthritis and osteoporosis, damage of cartilage and bone, and

CC formation of cartilage during embryonic development, for detecting the

CC physiological therapeutic repair of cartilage and muscle, for selection

CC and analysis, or for sorting, isolating or purification of chondrocytes

CC and muscle cells, for detecting regeneration of cartilage or chondrocytes

CC during transplantation of cartilage or chondrocytes during

CC transplantation of cartilage or chondrocytes, respectively, or of muscle

CC or muscle cells during transplantation of muscle or muscle cells,

CC respectively, and for studies of differentiation or chondrocytes or

CC muscle cells

XX Sequence 24 AA;

Query Match

Best Local Similarity 100.0%; Pred. No. 1.7e-15;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;  
KW developmental defect; inflammatory disease; dermatological; vulnary;  
KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

XX Rattus sp.

XX WO200190357-A1.

XX 29-NOV-2001.

XX 24-MAY-2001; 2001WO-NZ000099.

XX 24-MAY-2000; 2000US-0206650P.

XX 25-JUL-2000; 2000US-0221232P.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Watson JD, Strachan L, Sleeman M, Orrust R, Murison JG;  
XX Kumble KD;

XX WPI; 2002-122020/16.

XX New polynucleotides and polypeptides encoded by the polynucleotides  
PT isolated from skin cells, useful for treating skin wounds, cancers,  
PT growth and developmental defects, inflammatory diseases, or for  
PT modulating immune responses.

XX Example 2; Page 246; 466pp; English.

XX The present invention provides the protein and coding sequences of cDNAs  
CC isolated from human, murine and rat skin cell libraries. The sequences  
CC can be used in the development of therapeutic agents useful in the  
CC treatment of skin diseases, including skin wounds, cancer, growth  
CC defects, developmental defects and inflammatory diseases. The proteins  
CC have important roles in the induction of hair growth, cell proliferation  
CC and cell-cell interaction, in maintaining tissue integrity, in wound  
CC healing and in modulating immune responses. The present sequence is a  
CC polypeptide of the invention

XX Sequence 158 AA;

Query Match 1.8%; Score 21; DB 5; Length 158;  
Best Local Similarity 100.0%; Pred. No. 8.1e-12;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 RPVQINASLHPEPSKINIEH 656

DB 57 RPVQINASLHPEPSKINIEH 77

RESULT 43

AAAY76111

ID AAAY76111 standard; protein; 85 AA.

XX AC

XX AAAY76111;

XX 27-MAR-2000 (first entry)

XX Rat integrin homologue, SEQ ID NO:390.

XX Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;  
KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;  
KW secreted; transmembrane; inflammation; cancer; neurological disease;  
KW angiogenesis; tumour vascularisation; growth disorder;  
KW developmental disorder; skin wound; hair follicle disorder;  
KW anti-inflammatory; cytostatic; neuroprotective; vulnary.

XX Rattus sp.

XX WO9955865-A1.

XX 04-NOV-1999.

XX

PF 29-APR-1999; 99WO-NZ000051.  
XX 29-APR-1998; 98US-00069726.  
PR 09-NOV-1998; 98US-00188930.  
XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Strachan L, Sleeman M, Watson JD, Orrust R, Kumble A, Murison JG;  
XX WPI; 2000-072177/06.

XX N-PSDB; AA261819.

XX Novel polynucleotides useful for the treatment of various conditions  
PT including wounds and cancer.

XX Claim 4; Page 221; 235pp; English.

XX The invention relates to novel nucleic acid sequences derived from rat  
CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,  
CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying  
CC cells. Polypeptides of the invention may be used to treat inflammation,  
CC cancer and neurological diseases. The proteins may be used to stimulate  
CC the growth and motility of keratinocytes, to inhibit the growth of cancer  
CC cells, to modulate angiogenesis and tumour vascularisation, to modulate  
CC skin inflammation, to modulate epithelial cell growth and to inhibit  
CC binding of HIV-1 to leukocytes. The invention may also be used to treat  
CC growth and developmental defects, skin wounds and hair follicle  
CC disorders. Sequences AAAY75942-Y76123 represent polypeptides encoded by  
CC cDNA sequences derived from several mouse, rat or human skin cell types.  
CC Sequences AAAY75942-Y75947, AAAY76020-Y76021, AAAY76094-Y76104 and AAAY76119  
CC are proteins with an N-terminal signal sequence, indicating that they are  
CC secreted. Sequences AAAY75986-Y75989, AAAY76061-Y76071, AAAY76106-Y76109 and  
CC AAAY76121-Y76122 are proteins with one or more putative transmembrane  
CC domains

XX Sequence 85 AA;

Query Match 1.4%; Score 17; DB 3; Length 85;  
Best Local Similarity 100.0%; Pred. No. 4.4e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 NGTLKDSHSYQNAEFGS 544

DB 28 NGTLKDSHSYQNAEFGS 44

RESULT 44

AAAB56050

ID AAAB56050 standard; protein; 85 AA.

XX AC

XX AAAB56050;

XX 08-MAR-2001 (first entry)

XX Skin cell protein, SEQ ID NO: 390.

XX Rat; skin cell; cytostatic; antiinflammatory; anti-HIV; nootropic;  
KW neuroprotective; vulnary; immunomodulatory; vaccine; inhibition;  
KW keratinocyte growth stimulation; cancer; angiogenesis; inhibition;  
KW inflammation; neurological disease.

XX Rattus sp.

XX WO200069884-A2.

XX 23-NOV-2000.

XX 15-MAY-2000; 2000WO-NZ000075.

XX 14-MAY-1999; 99US-00312283.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX

PI Watson JD, Strachan L, Orrust R, Sleeman M, Kumble KD;  
PI Murison JG;  
XX WPI; 2001-007495/01.  
DR N-PSDB; AAC99752.  
XX  
XX New isolated polynucleotide used in the identification of genetic  
PT disorders and encoding polypeptides used for treating inflammatory  
PT disease, cancer and neurological diseases.  
XX  
PS Claim 4; Page 296; 352pp; English.  
XX  
XX The present sequence is a polypeptide which is expressed in mammalian  
CC skin cells. The polypeptide is useful for stimulating keratinocyte growth  
CC and motility, inhibiting the growth of cancer cells, modulating  
CC angiogenesis, inhibiting angiogenesis and vascularisation of tumours,  
CC modulating skin inflammation, stimulating the growth of epithelial cells,  
CC inhibiting the binding of human immunodeficiency virus (HIV)-1 to  
CC leukocytes, and treating inflammatory disease, cancer and neurological  
CC diseases. The polynucleotide can be used as a marker, in the  
CC identification of genetic disorders, and for the design of  
CC oligonucleotides for examining expression patterns  
XX  
SQ Sequence 85 AA;  
Query Match 1.4%; Score 17; DB 4; Length 85;  
Best Local Similarity 100.0%; Pred. No. 4.4e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 528 NGTLKDSHSYQNARFGS 544  
Db 28 NGTLKDSHSYQNARFGS 44  
RESULT 45  
ABBY72250  
ID ABB72250 standard; protein; 85 AA.  
XX AC ABB72250;  
XX  
DT 04-APR-2002 (first entry)  
XX  
DE Rat protein isolated from skin cells SEQ ID NO: 390.  
XX  
KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;  
KW developmental defect; inflammatory disease; dermatological; vulnary;  
KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.  
XX  
OS Rattus sp.  
XX  
XX WO200190357-A1.  
XX  
XX 29-NOV-2001.  
XX  
XX 24-MAY-2001; 2001WO-NZ000099.  
XX  
XX 24-MAY-2000; 2000US-0206550P.  
XX  
XX 25-JUL-2000; 2000US-0221232P.  
XX  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
XX Watson JD, Strachan L, Sleeman M, Orrust R, Murison JG;  
PI Kumble KD;  
XX  
XX WPI; 2002-122020/16.  
XX  
XX New polynucleotides and polypeptides encoded by the polynucleotides  
PT isolated from skin cells, useful for treating skin wounds, cancers,  
PT growth and developmental defects, inflammatory diseases, or for  
PT modulating immune responses.  
XX  
XX Example 2; Page 245; 466pp; English.  
XX

CC The present invention provides the protein and coding sequences of cDNAs  
CC isolated from human, murine and rat skin cell libraries. The sequences  
CC can be used in the development of therapeutic agents useful in the  
CC treatment of skin diseases, including skin wounds, cancer, growth  
CC defects, developmental defects and inflammatory diseases. The proteins  
CC have important roles in the induction of hair growth, cell proliferation  
CC and cell-cell interaction, in maintaining tissue integrity, in wound  
CC healing and in modulating immune responses. The present sequence is a  
XX polypeptide of the invention  
XX  
SQ Sequence 85 AA;  
Query Match 1.4%; Score 17; DB 5; Length 85;  
Best Local Similarity 100.0%; Pred. No. 4.4e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 528 NGTLKDSHSYQNARFGS 544  
Db 28 NGTLKDSHSYQNARFGS 44  
RESULT 46  
AAU19634  
ID AAU19634 standard; protein; 148 AA.  
XX  
XX AC AAU19634;  
XX  
DT 04-DEC-2001 (first entry)  
XX  
DE Human novel extracellular matrix protein, Seq ID No 284.  
XX  
XX Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;  
KW antianemic; antirheumatic; antisclerotic; cardiant; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;  
KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;  
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;  
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;  
KW Sezary syndrome; Gaucher's disease; neurological diseases;  
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
KW cardiac arrest; tachycardia; angina; infection; corneal infections;  
KW wound healing; immunogen; gene therapy; antisense; food additive.  
XX  
OS Homo sapiens.  
XX  
XX WO200155368-A1.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001348.  
XX  
XX 31-JAN-2000; 2000US-0179063P.  
XX  
XX 04-FEB-2000; 2000US-0180628P.  
XX  
XX 24-FEB-2000; 2000US-0184664P.  
XX  
XX 02-MAR-2000; 2000US-0186350P.  
XX  
XX 16-MAR-2000; 2000US-0189874P.  
XX  
XX 17-MAR-2000; 2000US-0190078P.  
XX  
XX 18-APR-2000; 2000US-0198123P.  
XX  
XX 19-MAY-2000; 2000US-0205515P.  
XX  
XX 07-JUN-2000; 2000US-0209467P.  
XX  
XX 28-JUN-2000; 2000US-0214886P.  
XX  
XX 30-JUN-2000; 2000US-0215135P.  
XX  
XX 07-JUL-2000; 2000US-0215647P.  
XX  
XX 07-JUL-2000; 2000US-0218880P.  
XX  
XX 11-JUL-2000; 2000US-0217496P.  
XX  
XX 14-JUL-2000; 2000US-0218290P.  
XX  
XX 26-JUL-2000; 2000US-0220963P.  
XX  
XX 14-AUG-2000; 2000US-0224518P.  
XX  
XX 14-AUG-2000; 2000US-0224519P.  
XX  
XX 14-AUG-2000; 2000US-0225213P.  
XX  
XX 14-AUG-2000; 2000US-0225214P.  
XX  
XX 14-AUG-2000; 2000US-0225266P.

PR	14-AUG-2000;	2000US-0225257P.
PR	14-AUG-2000;	2000US-0225256P.
PR	14-AUG-2000;	2000US-0325270P.
PR	14-AUG-2000;	2000US-0325447P.
PR	14-AUG-2000;	2000US-0325757P.
PR	14-AUG-2000;	2000US-0325758P.
PR	14-AUG-2000;	2000US-0325759P.
PR	14-AUG-2000;	2000US-0325627P.
PR	22-AUG-2000;	2000US-0326681P.
PR	22-AUG-2000;	2000US-0326686P.
PR	23-AUG-2000;	2000US-0327009P.
PR	30-SEP-2000;	2000US-0225892P.
PR	01-SEP-2000;	2000US-0225829P.
PR	01-SEP-2000;	2000US-0225934P.
PR	01-SEP-2000;	2000US-0225935P.
PR	05-SEP-2000;	2000US-0225950P.
PR	05-SEP-2000;	2000US-0225951P.
PR	06-SEP-2000;	2000US-0230437P.
PR	06-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231242P.
PR	08-SEP-2000;	2000US-0231243P.
PR	08-SEP-2000;	2000US-0231244P.
PR	08-SEP-2000;	2000US-0231413P.
PR	08-SEP-2000;	2000US-0231414P.
PR	08-SEP-2000;	2000US-0232080P.
PR	08-SEP-2000;	2000US-0232081P.
PR	12-SEP-2000;	2000US-0231968P.
PR	14-SEP-2000;	2000US-0232397P.
PR	14-SEP-2000;	2000US-0232398P.
PR	14-SEP-2000;	2000US-0232399P.
PR	14-SEP-2000;	2000US-0232400P.
PR	14-SEP-2000;	2000US-0232401P.
PR	14-SEP-2000;	2000US-0233054P.
PR	14-SEP-2000;	2000US-0233055P.
PR	21-SEP-2000;	2000US-0234223P.
PR	21-SEP-2000;	2000US-0234274P.
PR	25-SEP-2000;	2000US-0234997P.
PR	25-SEP-2000;	2000US-0234998P.
PR	26-SEP-2000;	2000US-0235484P.
PR	27-SEP-2000;	2000US-0235834P.
PR	27-SEP-2000;	2000US-0235836P.
PR	29-SEP-2000;	2000US-0236327P.
PR	29-SEP-2000;	2000US-0236367P.
PR	29-SEP-2000;	2000US-0236368P.
PR	29-SEP-2000;	2000US-0236369P.
PR	29-SEP-2000;	2000US-0236370P.
PR	02-OCT-2000;	2000US-0236803P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237039P.
PR	02-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239933P.
PR	13-OCT-2000;	2000US-0239937P.
PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241221P.
PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241803P.
PR	20-OCT-2000;	2000US-0241803P.
PR	01-NOV-2000;	2000US-0244617P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.

PR	08-NOV-2000;	2000US-0246526P
PR	08-NOV-2000;	2000US-0246527P
PR	08-NOV-2000;	2000US-0246528P
PR	08-NOV-2000;	2000US-0246532P
PR	08-NOV-2000;	2000US-02465609P
PR	08-NOV-2000;	2000US-0246610P
PR	08-NOV-2000;	2000US-0246611P
PR	08-NOV-2000;	2000US-0246613P
PR	17-NOV-2000;	2000US-0249207P
PR	17-NOV-2000;	2000US-0249208P
PR	17-NOV-2000;	2000US-0249209P
PR	17-NOV-2000;	2000US-0249210P
PR	17-NOV-2000;	2000US-0249211P
PR	17-NOV-2000;	2000US-0249212P
PR	17-NOV-2000;	2000US-0249213P
PR	17-NOV-2000;	2000US-0249214P
PR	17-NOV-2000;	2000US-0249215P
PR	17-NOV-2000;	2000US-0249216P
PR	17-NOV-2000;	2000US-0249217P
PR	17-NOV-2000;	2000US-0249218P
PR	17-NOV-2000;	2000US-0249244P
PR	17-NOV-2000;	2000US-0249245P
PR	17-NOV-2000;	2000US-0249264P
PR	17-NOV-2000;	2000US-0249265P
PR	17-NOV-2000;	2000US-0249297P
PR	17-NOV-2000;	2000US-0249299P
PR	01-DEC-2000;	2000US-0250160P
PR	01-DEC-2000;	2000US-0250391P
PR	05-DEC-2000;	2000US-0250300P
PR	05-DEC-2000;	2000US-0251988P
PR	05-DEC-2000;	2000US-0256719P
PR	06-DEC-2000;	2000US-0257147P
PR	08-DEC-2000;	2000US-0251856P
PR	08-DEC-2000;	2000US-0251868P
PR	08-DEC-2000;	2000US-0251869P
PR	08-DEC-2000;	2000US-0251989P
PR	08-DEC-2000;	2000US-0251990P
PR	11-DEC-2000;	2000US-0254097P
XX	JAN-2001;	2001US-0253678P.
XX	(HUMA-) HUMAN GENOME SCI INC.	
PI	Rosen CA, Barash SC, Ruben	
XX	WPI; 2001-465572/50.	
DR	N-PSDB; AAG31205.	
XX		
PT	Nucleic acid molecules encodi	
PT	proteins, used in preventing,	
PT	Alzheimer's and Parkinson's d	
XX		
XX	Claim 11; SEQ ID NO 284; 577p	
PS	The invention relates to isol	
CC	human secreted extracellular	
CC	and proteins are used to prev	
CC	humans, mice, rabbits, goats,	
CC	example, disorders associated	
CC	polynucleotide or a vector ex	
CC	diseases by gene therapy. Ant	
CC	regulate expression of Sps by	
CC	probing their expression	
CC	probes in diagnostic assays.	
CC	produce antibodies and to iden	
CC	of the Sps. The anti-(Ssp) ant	
CC	down regulate expression and	
CC	detecting the presence of Sps	
CC	example; immune/autoimmune di	
CC	virus) infections, anaemia, r	
CC	cancers and hyperproliferativ	
CC	the breast or liver, Szary s	
CC	diseases (e.g. Alzheimer's di	

CC /cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina),  
CC infections caused by bacteria, viruses and fungi and ocular disorders  
CC (e.g. corneal infections). Other uses include wound healing, maintenance  
CC of organs before transplantation, support of cell culture of primary

Query Match 1.38; Score 15; DB 4; Length 148;  
Best Local Similarity 100.0%; Pred. NO. 6.9e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGDSNIPW 178  
Db 47 DIVVLGDSNIPW 61

## RESULT 47

AAU19794  
ID AAU19794 standard; protein: 148 AA.

XX AAU19794;

DT 04-DEC-2001 (first entry)

DE Human novel extracellular matrix protein, Seq ID No 444.

XX Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;  
KW antianemic; antirheumatic; antisclerotic; cardiac; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; HIV infection; anaemia;  
KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;  
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;  
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;  
KW Sezary syndrome; Gaucher's disease; neurological diseases;  
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
KW cardiac arrest; tachycardia; angina; infection; corneal infections;  
KW wound healing; immunogen; gene therapy; antisense; food additive.

XX Homo sapiens.

XX WO200155368-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001348.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184654P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198223P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 30-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232089P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234988P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241828P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.

[illegible]

```

PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244817P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2002-470713/50.
XX N-PSDB; ABQ66529.
XX
PT New nucleic acid encoding human proteins, useful for diagnosis, treatment
PT and prevention of e.g. osteoporosis, also related polypeptides and
PT antibodies.
XX
PS Claim 11; SEQ ID NO 284; 235pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABQ66521-ABQ66785) and proteins
CC (ABP47846-ABP4810) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html?DocID=99909764870
XX
SQ Sequence 148 AA;
Query Match 1.3%; Score 15; DB 5; Length 148;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 164 DIVIVLDGNSIYPW 178
Db 47 DIVIVLDGNSIYPW 61
RESULT 49
ABP48014
ID ABP48014 standard; protein; 148 AA.
XX
AC ABP48014;
XX
XX 23-AUG-2002 (first entry)
DE Human polypeptide SEQ ID NO 444.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnary;
KW antiparkinsonian; antiskinkling; antianemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX

```

```

OS Homo sapiens.
XX US2002042386-A1.
XX
PD 11-APR-2002.
XX
PF 17-JAN-2001; 2001US-00764870.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226888P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2002-470713/50.
DR N-PSDB; ABQ66589.
XX
PT New nucleic acid encoding human proteins, useful for diagnosis, treatment
PT and prevention of e.g. osteoporosis, also related polypeptides and
PT antibodies.
XX
PS Claim 11; SEQ ID NO 444; 235pp + Sequence Listing; English.

```

XX The invention relates to novel genes (ABQ66521-ABQ66785) and proteins  
 CC (ABQ47846-ABQ48110) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from USPTO at seqdata.uspto.gov/sequence.html?docid=999909764870  
 XX  
 SQ Sequence 148 AA;

Query Match 1.3%; Score 15; DB 5; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178  
 Db |||||  
 47 DIVIVLDGNSIYPW 61

RESULT 50  
 ADC10816  
 ID ADC10816 standard; protein; 148 AA.

XX AC ADC10816;  
 XX DT 18-DEC-2003 (first entry)  
 XX DE Human extracellular matrix protein from gene 9.

XX Extracellular matrix protein; cytostatic; antibacterial; virucide;  
 KW neuroprotective; gynaecological; gastrointestinal; cardiac;  
 KW cardiovascular; nephrotropic; antiinflammatory; muscular; Gen;  
 KW respiratory; Gen; immunosuppressive; cerebroprotective; viral infection;  
 KW neoplastic; antiallergic; cancer; bacterial infection; viral infection;  
 KW neural disorder; immune system disorder; blood disorder;  
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
 KW pulmonary disorder; cardiovascular disorder; renal disorder;  
 KW inflammatory disorder; proliferative disorder; Human.

XX OS Homo sapiens.  
 XX PN US2003059875-A1.  
 XX PD 27-MAR-2003.  
 XX PF 19-APR-2002; 2002US-00125540.

XX PR 31-JAN-2000; 2000US-0179063P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190078P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 11-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226686P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239335P.  
 PR 13-OCT-2000; 2000US-0239337P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244461P.

PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 06-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0253678P.  
 PR 17-JAN-2001; 2001US-00764870.  
 (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Ruben SM, Barash SC;

PI WPI; 2003-743765/70.

XX N-PSDB; ADC10551.

XX New isolated nucleic acids and polypeptides, useful for diagnosing, treating, and/or preventing disorders, such as cancer, infections, cardiovascular and inflammatory diseases.

PS Claim 11; SEQ ID NO 284; 235pp; English.

XX The invention relates to an isolated nucleic acid molecule (cDNA) encoding a human extracellular matrix protein, representing one of 161 novel genes. Also included are recombinant vectors, host cells (expressing the protein), the extracellular matrix proteins (including their fragments, epitopes and homologues), an isolated antibody that binds specifically to the protein, diagnosing a pathological condition or susceptibility to a pathological condition (comprising determining the presence or absence of a mutation in the nucleic acid and diagnosing a condition based on the presence or absence of the mutation), diagnosing a pathological condition or susceptibility to a pathological condition

CC (comprising determining the presence or amount of expression of the protein in a biological sample and diagnosing a condition based on the presence or amount of expression of the protein), preventing, treating or ameliorating a medical condition by administering the nucleic acid or protein to a mammalian subject, identifying a binding partner to the protein, the gene corresponding to the cDNA sequence, and identifying an activity in a biological assay (comprising expressing the nucleic acid in a cell, isolating the supernatant, detecting an activity in a biological assay and identifying the protein in the supernatant having the activity). The nucleic acids and proteins display the following activities: Cytostatic, antibacterial, Virucide, Neuroprotective, Gynaecological, Gastrointestinal-Gen, Cardiant, Cardiovascular-Gen, Nephrotropic, Antiinflammatory, Muscular-Gen, Respiratory-Gen,

Query Match 1.3%; Score 15; DB 7; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178  
 |||||  
 DB 47 DIVIVLDGNSIYPW 61

RESULT 51

ADC10976  
 ID ADC10976 standard; protein; 148 AA.

XX AC ADC10976;

XX DT 18-DEC-2003 (first entry)

XX DE Human protein from extracellular matrix gene 9 #2.

XX Extracellular matrix protein; cytostatic; antibacterial; virucide; neuroprotective; gynaecological; gastrointestinal-Gen; cardiant; cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen; respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic; neotropic; antiallergic; cancer; bacterial infection; viral infection; neural disorder; immune system disorder; blood disorder; muscular disorder; reproductive disorder; gastrointestinal disorder; pulmonary disorder; cardiovascular disorder; renal disorder; inflammatory disorder; proliferative disorder; Human.

XX Homo sapiens.

XX US2003059875-A1.

XX PD 27-MAR-2003.

XX PF 19-APR-2002; 2002US-00125540.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 11-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218299P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225213P.

XX PR 14-AUG-2000; 2000US-0225214P.







```

PT fibrosis.
PS Example 24; Page 93-94; 248pp; English.
XX
XX
CC The present invention relates to novel antibodies that specifically bind
CC to very late activation (VLA-1; beta1 containing integrins) antigens and
CC methods of using these antibodies to treat immunological disorders. The
CC anti-VLA-1 antibodies are useful for preventing or treating VLA-1-
CC mediated immunological or inflammatory disorders such as skin related
CC conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal
CC proliferation of hair follicle cells), fibrosis (e.g. kidney or lung
CC fibrosis), allergic rhinitis, respiratory distress syndrome, asthma,
CC bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-
CC intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease,
CC gastritis, irritable bowel syndrome, colitis and colorectal cancer),
CC vascular diseases (e.g. atherosclerosis), thyroiditis, aplastic anaemia,
CC periarthritis nodosa, Hodgkin's disease, rheumatic fever, osteoarthritis,
CC autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid
CC arthritis, systemic lupus erythematosus and multiple sclerosis), renal
CC failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis,
CC polynovositis, hypersensitivity (e.g. delayed type hypersensitivity or
CC immediate hypersensitivity), graft and transplant rejections, graft
CC versus host disease, conjunctivitis, swelling occurring after injury,
CC myocardial ischaemia or endotoxin shock syndrome. The present invention
CC rat 1-I domain protein. This sequence is used in the exemplification of
CC the invention
XX
SQ Sequence 192 AA;

Query Match 1.3%; Score 15; DB 6; Length 192;
Best Local Similarity 100.0%; Pred. No. 8.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLGGNSIYPW 178
Db |||||
4 DIVIVLGGNSIYPW 18

RESULT 55
AAU76851
ID AAU76851 standard; protein; 195 AA.
XX
XX AAU76851;
XX
XX 21-MAY-2002 (first entry)
XX
XX Human integrin alpha subunit Alpha 1 (CD49a) A domain.
XX
XX Human; integrin alpha subunit; A domain; CD49a; integrin beta subunit;
XX A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
XX ischaemia-reperfusion injury; immune complex; parasitic disease; Alpha 1;
XX antinflammatory; vasotropic; antiparasitic; vulnery; gene therapy.
XX
XX Homo sapiens.
XX
XX WO200209737-A1.
XX
XX 07-FEB-2002.
XX
XX 31-JUL-2001; 2001WO-US023957.
XX
XX 31-JUL-2000; 2000US-0221950P.
XX
XX 11-JAN-2001; 2001US-00758493.
XX
XX 13-MAR-2001; 2001US-00805354.

```

```

XX PA (GEO ) GEN HOSPITAL CORP.
XX PI
XX PI Annaout AM, Li R, Xiong J;
XX DR WPI; 2002-188687/24.
XX
XX PT Novel high affinity integrin polypeptide useful for treating restenosis
XX PT and parasitic diseases, comprises all or part of variant integrin alpha
XX PT subunit A domain or variant integrin beta subunit A-like domain.
XX
XX PS Example 2; Fig 5; 55pp; English.
XX
XX CC The invention relates to a high affinity integrin polypeptide comprising
XX CC all or part of a variant integrin alpha subunit A domain or a variant
XX CC integrin beta subunit A-like domain. The polypeptide, preferably the
XX CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
XX CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
XX CC V at residue 315 and A at residue 320 have been replaced by C, is useful
XX CC for determining if a test compound is a candidate compound for binding to
XX CC CD11b or for treating an inflammatory disorder, by contacting a test
XX CC compound with the polypeptide and determining if the test compound binds
XX CC to the polypeptide. The integrin subunits are useful for reducing
XX CC skeletal muscle injury, for treating disorders caused by ischaemia-
XX CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
XX CC to purify variant integrin polypeptide ligands and as bait proteins in
XX CC two-hybrid or three-hybrid assays. This sequence represents the human
XX CC integrin alpha subunit Alpha 1 (CD49a) A domain
XX
XX SQ Sequence 195 AA;
Query Match 1.3%; Score 15; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 164 DIVIVLDGNSIYPW 178
Db 6 DIVIVLDGNSIYPW 20
RESULT 56
AAU76860
ID AAU76860 standard; protein; 195 AA.
XX
XX AC AAU76860;
XX
XX DT 21-MAY-2002 (first entry)
XX
XX DE Human integrin alpha subunit Alpha 1 (CD49a) variant A domain.
XX
XX KW Human; integrin alpha subunit; A domain; CD49a; integrin beta subunit;
XX KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
XX KW ischaemia-reperfusion injury; immune complex; parasitic disease; mutant;
XX KW anti-inflammatory; vasotropic; antiparasitic; vulnerary; gene therapy;
XX KW mutein; Alpha 1.
XX
XX OS Homo sapiens.
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 193
XX FT /note= "Wild-type Ile substituted by any other amino
XX FT acid"
XX
XX PN WO200209737-A1.
XX
XX PD 07-FEB-2002.
XX
XX PP 31-JUL-2001; 2001WO-US023957.
XX
XX PR 31-JUL-2000; 2000US-0221950P.
XX PR 11-JAN-2001; 2001US-00758493.
XX PR 13-MAR-2001; 2001US-00805354.

```

```

XX PA (GEO ) GEN HOSPITAL CORP.
XX PI
XX PI Annaout AM, Li R, Xiong J;
XX DR WPI; 2002-188687/24.
XX
XX PT Novel high affinity integrin polypeptide useful for treating restenosis
XX PT and parasitic diseases, comprises all or part of variant integrin alpha
XX PT subunit A domain or variant integrin beta subunit A-like domain.
XX
XX PS Claim 45; Page; 55pp; English.
XX
XX CC The invention relates to a high affinity integrin polypeptide comprising
XX CC all or part of a variant integrin alpha subunit A domain or a variant
XX CC integrin beta subunit A-like domain. The polypeptide, preferably the
XX CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
XX CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
XX CC V at residue 315 and A at residue 320 have been replaced by C, is useful
XX CC for determining if a test compound is a candidate compound for binding to
XX CC CD11b or for treating an inflammatory disorder, by contacting a test
XX CC compound with the polypeptide and determining if the test compound binds
XX CC to the polypeptide. The integrin subunits are useful for reducing
XX CC skeletal muscle injury, for treating disorders caused by ischaemia-
XX CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
XX CC to purify variant integrin polypeptide ligands and as bait proteins in
XX CC two-hybrid or three-hybrid assays. This sequence represents a human
XX CC integrin alpha subunit Alpha 1 (CD49a) variant A domain. Note: This
XX CC variant sequence is not featured in the specification but has been
XX CC derived from the wild-type protein shown in AAU76851
XX
XX SQ Sequence 195 AA;
Query Match 1.3%; Score 15; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 164 DIVIVLDGNSIYPW 178
Db 6 DIVIVLDGNSIYPW 20
RESULT 57
AAB50041
ID AAB50041 standard; protein; 214 AA.
XX
XX AC AAB50041;
XX
XX DT 14-MAR-2001 (first entry)
XX
XX DE Rat alpha1 integrin domain.
XX
XX KW Rat; alpha1 integrin domain; alphasbeta1 function blocking antibody;
XX KW inflammatory disorder; rheumatoid arthritis; skin condition; asthma;
XX KW bronchitis; headache; antipyretic; fever; gastrointestinal;
XX KW vascular disease; autoimmune; respiratory distress syndrome;
XX KW endotoxin shock syndrome; atherosclerosis.
XX
XX OS Rattus sp.
XX
XX PN WO2000072881-A1.
XX
XX PD 07-DEC-2000.
XX
XX PP 01-JUN-2000; 2000WO-US015004.
XX
XX PR 01-JUN-1999; 99US-0137038P.
XX PR 29-FEB-2000; 2000US-0185336P.
XX
XX PA (BIO ) BIOGEN INC.
XX
XX PI De Fougereolles A, Gotwals P, Lobb R, Koteliensky V;
XX

```

DR WPI; 2001-061448/07.  
 XX Use of blocking monoclonal antibody capable of binding to an epitope of  
 PT VLA-1 for treating inflammatory disorders, in particular arthritis.  
 PS Example 14; Fig 14; 60pp; English.  
 XX The present invention relates to the use of an alphanbetal function  
 CC blocking antibody capable of binding an epitope of human alphan integrin  
 CC domain (VLA-1; see AAB50043). The antibody of the present invention is  
 CC useful for treating an inflammatory disorder e.g. rheumatoid arthritis,  
 CC skin related conditions such as psoriasis, eczema, burns and dermatitis,  
 CC asthma, bronchitis, menstrual cramps, tendinitis, bursitis, and the  
 CC treatment of pain and headaches, or as an antipyretic, for the treatment  
 CC of fever, gastrointestinal conditions such as inflammatory bowel  
 CC diseases, Crohn's disease, gastritis and vascular diseases, migraine  
 CC headaches, periarthritis nodosa, thyroiditis, aplastic anaemia, Hodgkin's  
 CC disease, rheumatic fever, type I diabetes, myasthenia gravis, multiple  
 CC sclerosis, sarcoidosis, nephrotic syndrome, myocardial ischaemia,  
 CC allergic rhinitis, respiratory distress syndrome, endotoxin shock  
 CC syndrome and atherosclerosis. The present sequence is the rat alphan  
 CC integrin domain  
 XX Sequence 214 AA;  
 SQ

Query Match 1.3%; Score 15; DB 4; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLGGNSIYPW 178  
 |||||  
 DB 22 DIVIVLGGNSIYPW 36

RESULT 58  
 AAB50042  
 ID AAB50042 standard; protein; 214 AA.  
 AC AAB50042;  
 XX  
 DT 14-MAR-2001 (first entry)  
 DE Human alphan integrin domain.  
 XX Human; alphan integrin domain; alphanbetal function blocking antibody;  
 KW inflammatory disorder; rheumatoid arthritis; skin condition; asthma;  
 KW bronchitis; headache; antipyretic; fever; gastrointestinal;  
 KW vascular disease; autoimmune; respiratory distress syndrome;  
 KW endotoxin shock syndrome; atherosclerosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200072881-A1.  
 XX  
 PD 07-DEC-2000.  
 XX  
 PF 01-JUN-2000; 2000WO-US015004.  
 XX  
 PR 01-JUN-1999; 99US-0137038P.  
 PR 29-FEB-2000; 2000US-0185336P.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 XX  
 PI De Fougereolles A, Gotwals P, Lobb R, Koteliensky V;  
 XX  
 DR WPI; 2001-061448/07.  
 XX  
 XX Use of blocking monoclonal antibody capable of binding to an epitope of  
 PT VLA-1 for treating inflammatory disorders, in particular arthritis.  
 XX Example 14; Fig 15; 60pp; English.  
 XX The present invention relates to the use of an alphanbetal function

CC blocking antibody capable of binding an epitope of human alphan integrin  
 CC domain (VLA-1; see AAB50043). The antibody of the present invention is  
 CC useful for treating an inflammatory disorder e.g. rheumatoid arthritis,  
 CC skin related conditions such as psoriasis, eczema, burns and dermatitis,  
 CC asthma, bronchitis, menstrual cramps, tendinitis, bursitis, and the  
 CC treatment of pain and headaches, or as an antipyretic, for the treatment  
 CC of fever, gastrointestinal conditions such as inflammatory bowel  
 CC diseases, Crohn's disease, gastritis and vascular diseases, migraine  
 CC headaches, periarthritis nodosa, thyroiditis, aplastic anaemia, Hodgkin's  
 CC disease, rheumatic fever, type I diabetes, myasthenia gravis, multiple  
 CC sclerosis, sarcoidosis, nephrotic syndrome, myocardial ischaemia,  
 CC allergic rhinitis, respiratory distress syndrome, endotoxin shock  
 CC syndrome and atherosclerosis. The present sequence is the human alphan  
 CC integrin domain  
 XX Sequence 214 AA;  
 SQ

Query Match 1.3%; Score 15; DB 4; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLGGNSIYPW 178  
 |||||  
 DB 22 DIVIVLGGNSIYPW 36

RESULT 59  
 AAB53539  
 ID AAB53539 standard; protein; 214 AA.  
 AC AAB53539;  
 XX  
 DT 02-APR-2003 (first entry)  
 DE Rat alpha 1-I domain protein #2.  
 XX  
 KW Rat; very late activation antigen; VLA-1; betal containing integrin;  
 KW immunological disorder; inflammatory disorder; skin related condition;  
 KW psoriasis; eczema; burn; dermatitis; respiratory distress syndrome;  
 KW fibrosis; allergic rhinitis; asthma; bronchitis; tendinitis; bursitis;  
 KW fever; migraine headache; inflammatory bowel disease; Crohn's disease;  
 KW irritable bowel syndrome; colitis; colorectal cancer; vascular disease;  
 KW atherosclerosis; thyroiditis; aplastic anaemia; periarthritis nodosa;  
 KW gastritis; Hodgkin's disease; rheumatic fever; autoimmune disease;  
 KW osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis;  
 KW systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome;  
 KW renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis;  
 KW hypersensitivity; graft rejection; transplant rejection; conjunctivitis;  
 KW graft versus host disease; myocardial ischaemia; alpha 1-I protein.  
 XX  
 OS Rattus sp.  
 XX  
 PN WO200283854-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 12-APR-2002; 2002WO-US011521.  
 XX  
 PR 13-APR-2001; 2001US-0283794P.  
 PR 06-JUL-2001; 2001US-0303689P.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 XX  
 PI Lyne PD, Garber EA, Saldanha JW, Karpusas M;  
 XX  
 DR WPI; 2003-093009/08.  
 XX  
 XX New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-  
 PT mediated immunological or inflammatory disorders, e.g. psoriasis, eczema,  
 PT burns, dermatitis, and abnormal proliferation of hair follicle cells or  
 PT fibrosis.  
 XX Example 18; Fig 11A; 248pp; English.

XX The present invention relates to novel antibodies that specifically bind  
 CC to very late activation (VLA-1; beta1 containing integrins) antigens and  
 CC methods of using these antibodies to treat immunological disorders. The  
 CC anti-VLA-1 antibodies are useful for preventing or treating VLA-1-  
 CC mediated immunological or inflammatory disorders such as skin related  
 CC conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal  
 CC proliferation of hair follicle cells), fibrosis (e.g. kidney or lung  
 CC fibrosis), allergic rhinitis, respiratory distress syndrome, asthma,  
 CC bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-  
 CC intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease,  
 CC gastritis, irritable bowel syndrome, colitis and colorectal cancer),  
 CC vascular diseases (e.g. atherosclerosis), thyroiditis, aplastic anaemia,  
 CC periarthritis nodosa, Hodgkin's disease, rheumatic fever, osteoarthritis,  
 CC autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid  
 CC arthritis, systemic lupus erythematosus and multiple sclerosis), renal  
 CC failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis,  
 CC polymyositis, hypersensitivity (e.g. delayed type hypersensitivity or  
 CC immediate hypersensitivity), graft and transplant rejections, graft  
 CC versus host disease, conjunctivitis, swelling occurring after injury,  
 CC myocardial ischaemia or endotoxin shock syndrome. The present sequence is  
 CC rat 1-I domain protein. This sequence is used in the exemplification of  
 CC the invention  
 CC  
 CC SQ Sequence 214 AA;

Query Match 1.3%; Score 15; DB 6; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 164 DIVVLGSGNSIYPW 178  
 DB 22 DIVVLGSGNSIYPW 36  
 |||||

RESULT 60  
 AAE33540  
 ID AAE33540 standard; protein; 214 AA.  
 XX  
 AC AAE33540;  
 XX  
 DT 02-APR-2003 (first entry)  
 DE Human alpha 1-I domain protein #2.  
 KW Human; very late activation antigen; VLA-1; beta1 containing integrin;  
 KW immunological disorder; inflammatory disorder; skin related condition;  
 KW psoriasis; eczema; burn; dermatitis; respiratory distress syndrome;  
 KW fibrosis; allergic rhinitis; asthma; bronchitis; tendonitis; bursitis;  
 KW fever; migraine headache; inflammatory bowel disease; Crohn's disease;  
 KW irritable bowel syndrome; colitis; colorectal cancer; vascular disease;  
 KW atherosclerosis; thyroiditis; aplastic anaemia; periarthritis nodosa;  
 KW gastritis; Hodgkin's disease; rheumatic fever; autoimmune disease;  
 KW osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis;  
 KW systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome;  
 KW renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis;  
 KW hypersensitivity; graft rejection; transplant rejection; conjunctivitis;  
 KW graft versus host disease; myocardial ischaemia; alpha 1-I protein.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283854-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 XX 12-APR-2002; 2002WO-US011521.  
 PF  
 XX 13-APR-2001; 2001US-0283794P.  
 PR  
 PR 06-JUL-2001; 2001US-0303689P.  
 XX  
 XX (BIOJ ) BIOGEN INC.  
 PA  
 XX Lyne PD, Garber EA, Saldanha JW, Karpusas M;  
 PI

XX WPI; 2003-093009/08.  
 XX New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-  
 PT mediated immunological or inflammatory disorders, e.g. psoriasis, eczema,  
 PT burns, dermatitis, and abnormal proliferation of hair follicle cells or  
 PT fibrosis.  
 XX Example 18; Fig 12; 248pp; English.  
 PS  
 CC The present invention relates to novel antibodies that specifically bind  
 CC to very late activation (VLA-1; beta1 containing integrins) antigens and  
 CC methods of using these antibodies to treat immunological disorders. The  
 CC anti-VLA-1 antibodies are useful for preventing or treating VLA-1-  
 CC mediated immunological or inflammatory disorders such as skin related  
 CC conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal  
 CC proliferation of hair follicle cells), fibrosis (e.g. kidney or lung  
 CC fibrosis), allergic rhinitis, respiratory distress syndrome, asthma,  
 CC bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-  
 CC intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease,  
 CC gastritis, irritable bowel syndrome, colitis and colorectal cancer),  
 CC vascular diseases (e.g. atherosclerosis), thyroiditis, aplastic anaemia,  
 CC periarthritis nodosa, Hodgkin's disease, rheumatic fever, osteoarthritis,  
 CC autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid  
 CC arthritis, systemic lupus erythematosus and multiple sclerosis), renal  
 CC failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis,  
 CC polymyositis, hypersensitivity (e.g. delayed type hypersensitivity or  
 CC immediate hypersensitivity), graft and transplant rejections, graft  
 CC versus host disease, conjunctivitis, swelling occurring after injury,  
 CC myocardial ischaemia or endotoxin shock syndrome. The present sequence is  
 CC human 1-I domain protein. This sequence is used in the exemplification of  
 CC the invention  
 CC  
 CC SQ Sequence 214 AA;

Query Match 1.3%; Score 15; DB 6; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-06;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 164 DIVVLGSGNSIYPW 178  
 DB 22 DIVVLGSGNSIYPW 36  
 |||||

RESULT 61  
 ADE86652  
 ID ADE86652 standard; protein; 1151 AA.  
 XX  
 AC ADE86652;  
 XX  
 DT 29-JAN-2004 (first entry)  
 DE Human integrin alpha 1 subunit.  
 KW human; secreted protein; cancer; liver disorder; hepatitis;  
 KW neural disorder; Alzheimer's disease; integrin alpha 1 subunit.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003129685-A1.  
 XX  
 PD 10-JUL-2003.  
 XX  
 XX 18-APR-2001; 2001US-00836353.  
 PF  
 XX 28-OCT-1998; 98US-0105971P.  
 PR  
 PR 27-OCT-1999; 99WO-US025031.  
 PR  
 PR 19-APR-2000; 2000US-0198407P.  
 XX  
 XX (NIJ/) NI J.  
 PA (YOUNG) YOUNG P E.  
 PA (KENN) KENNY J J.  
 PA (OLSE) OLSEN H S.



CC therapeutic, diagnostic and research methods. The polynucleotides may be  
 CC used in the field of molecular biology as hybridisation probes, primers  
 CC for PCR, for chromosome and gene mapping, for the recombinant production  
 CC of protein, or in generation of anti-sense DNA or RNA. The  
 CC polynucleotides are useful in diagnostics as expressed sequence tags  
 CC (ESTs) for identifying expressed genes or for physical mapping of the  
 CC human genome. The proteins may be used as molecular weight markers, or as  
 CC nutritional sources or supplements. The proteins may be used to maintain  
 CC and expand cell population in a totipotent or pluripotent state  
 CC useful for re-engineering damaged or diseased tissues, transplantation.  
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The  
 CC polynucleotides and proteins are useful for preventing, treating or  
 CC ameliorating disorders involving aberrant protein expression or  
 CC biological activity, e.g. haematopoietic disorders, central/peripheral  
 CC nervous system diseases, mechanical and traumatic disorders, non-healing  
 CC wounds, immune deficiencies and disorders, infectious diseases caused by  
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic  
 CC reactions and conditions, coagulation disorders, or cancer. The  
 CC polynucleotide sequences of the invention were assembled from ESTs  
 CC isolated mainly by sequencing by hybridisation, and in some cases,  
 CC sequences obtained from one or more public databases. Note: The sequence  
 CC data for this patent did not form part of the printed specification, but  
 CC was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1179 AA;

Query Match 1.3%; Score 15; DB 5; Length 1179;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178  
 Db 172 DIVIVLDGNSIYPW 186  
 |||||

RESULT 64  
 ABUS4466  
 ID ABUS4466 standard; protein; 1179 AA.  
 AC ABUS4466;  
 XX  
 DT 12-MAR-2003 (first entry)  
 DE Human tumour endothelial marker TEM 30.  
 XX  
 KW Human; endothelial cell; EC; tumour endothelial cell; TEM; NEM;  
 KW Tumour endothelial marker; normal endothelial marker; PEM;  
 KW pan-endothelial marker; polycystic kidney disease; psoriasis;  
 KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;  
 KW neoangiogenesis; immune response; cytostatic; antidiabetic;  
 KW ophthalmological; antirheumatic; antiarthritic; antipsoriatic.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200283874-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 10-APR-2002; 2002WO-US08253.  
 XX  
 PR 11-APR-2001; 2001US-0282850P.  
 PR 06-FEB-2002; 2002US-0354262P.  
 XX  
 PA (UYJO ) UNIV JOHNS HOPKINS.  
 XX  
 PI Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;  
 XX  
 DR WPI; 2003-093016/08.  
 DR N-PSDB; AEX72038.

PT New purified human transmembrane protein, designated as tumor endothelial  
 marker (TEM) 3, useful for detecting, diagnosing or treating tumors,

PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or  
 PT psoriasis.  
 XX  
 PS Disclosure; Page 267-269; 374pp; English.  
 XX  
 CC The present invention relates to a novel method for the isolation of  
 CC endothelial cells (ECs), and the identification of genes expressed in  
 CC normal and tumour ECs. Tumour endothelial marker (TEM), normal  
 CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are  
 CC identified in human ECs. The human EC marker proteins and the  
 CC polynucleotide sequences encoding them are useful for detecting,  
 CC diagnosing or treating tumours as well as polycystic kidney disease,  
 CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also  
 CC useful for inhibiting neoangiogenesis or tumour angiogenesis, for  
 CC inducing an immune response to tumour endothelial cells in a patient, or  
 CC for identifying candidate drugs for treating tumours. The present  
 CC sequence represents a human TEM or NEM protein of the invention  
 XX  
 SQ Sequence 1179 AA;

Query Match 1.3%; Score 15; DB 6; Length 1179;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178  
 Db 172 DIVIVLDGNSIYPW 186  
 |||||

RESULT 65  
 ABB90788  
 ID ABB90788 standard; protein; 1180 AA.

AC ABB90788;  
 XX  
 DT 30-MAY-2002 (first entry)  
 DE Rat Tumour Endothelial Marker polypeptide SEQ ID NO 307.  
 XX  
 KW Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;  
 KW normal endothelial marker; pan-endothelial marker; immunostimulant;  
 KW antiangiogenic; tumour; neoangiogenesis; vascularised tumour;  
 KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;  
 KW psoriasis.  
 XX  
 OS Rattus sp.  
 XX  
 FN WO200210217-A2.  
 XX  
 PD 07-FEB-2002.  
 XX  
 PF 01-AUG-2001; 2001WO-US024031.  
 XX  
 PR 02-AUG-2000; 2000US-0222599P.  
 PR 11-AUG-2000; 2000US-0224360P.  
 PR 11-APR-2001; 2001US-0282850P.  
 XX  
 PA (UYJO ) UNIV JOHNS HOPKINS.  
 XX  
 PI St Croix B, Kinzler KW, Vogelstein B;  
 XX  
 DR WPI; 2002-291856/33.  
 DR N-PSDB; ABL92141.

PT An isolated molecule comprising an antibody variable region which  
 specifically binds to an extracellular domain of a tumor endothelial  
 marker (TEM) protein, useful for inhibiting tumor growth.

PS Disclosure; Page 314-318; 331pp; English.

CC The invention relates to an isolated molecule comprising an antibody  
 variable region which specifically binds to an extracellular domain of a  
 tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,



CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM  
 CC proteins have cytostatic, immunostimulant and antiangiogenic activity.  
 CC They are useful for inhibiting tumour growth, neoangiogenesis in subjects  
 CC bearing a vascularised tumour, polycystic kidney disease, diabetic  
 CC retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM  
 CC genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789)  
 CC are disclosed, as are marker oligonucleotide sequences: tumour  
 CC endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal  
 CC endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers  
 CC (PEM) ABL91903-ABL91995

XX SQ Sequence 1180 AA;

Query Match 1.3%; Score 15; DB 5; Length 1180;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVIVLDGNSNIYPW 178  
 Db 172 DIVIVLDGNSNIYPW 186  
 |||||

## RESULT 66

ABUS4495  
 ID ABUS4495 standard; protein; 1180 AA.

XX AC ABUS4495;

XX DT 12-MAR-2003 (first entry)

XX DE Mouse tumour endothelial marker mTEM 30.

XX KW Mouse; endothelial cell; EC; tumour endothelial cell; TEM; NEM;  
 KW Tumour endothelial marker; normal endothelial marker; PEM;  
 KW pan-endothelial marker; polycystic kidney disease; psoriasis;  
 KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;  
 KW neoangiogenesis; immune response; cytostatic; antidiabetic;  
 KW ophthalmological; antirheumatic; antiarthritic; antipsoriatic.

XX OS Mus sp.

XX PN WO200283874-A2.

XX XX 24-OCT-2002.

XX PF 10-APR-2002; 2002WO-US008253.

XX PR 11-APR-2001; 2001US-0282850P.

XX PR 06-FEB-2002; 2002US-0354262P.

XX XX (UUYO ) UNIV JOHNS HOPKINS.

XX PI Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;

XX DR WPI; 2003-093016/08.

XX DR N-PSDB; ABX72066.

XX PT New purified human transmembrane protein, designated as tumor endothelial  
 PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors,  
 PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or  
 PT psoriasis.

XX PS Disclosure; Page 356-359; 374pp; English.

XX CC The present invention relates to a novel method for the isolation of  
 CC endothelial cells (ECs), and the identification of genes expressed in  
 CC normal and tumour ECs. Tumour endothelial marker (TEM), normal  
 CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are  
 CC identified in human ECs. The human EC marker proteins and the  
 CC polynucleotide sequences encoding them are useful for detecting,  
 CC diagnosing or treating tumours as well as polycystic kidney disease,  
 CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also  
 CC useful for inhibiting neoangiogenesis or tumour angiogenesis, for

CC inducing an immune response to tumour endothelial cells in a patient, or  
 CC for identifying candidate drugs for treating tumours. The present  
 CC sequence represents a mouse TEM protein

XX SQ Sequence 1180 AA;

Query Match 1.3%; Score 15; DB 6; Length 1180;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVIVLDGNSNIYPW 178  
 Db 172 DIVIVLDGNSNIYPW 186  
 |||||

## RESULT 67

ADE63568  
 ID ADE63568 standard; protein; 1180 AA.

XX AC ADE63568;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein P18614, SEQ ID NO 9512.

XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX PN WO2003016475-A2.

XX XX 27-FEB-2003.

XX XX 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX PA (GEHO ) GEN HOSPITAL CORP.

XX PA (FARB ) BAYER AG.

XX PI Woolf C, D'urso D, Befort K, Costigan M;

XX DR WPI; 2003-288312/26.

XX DR GENBANK; P18614.

XX PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating

CC pain (e.g. spinal segmental nerve injury (Chung)), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1180 AA;

Query Match 1.3%; Score 15; DB 7; Length 1180;

Best Local Similarity 100.0%; Pred. No. 4.2e-05; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 0;

QY 164 DIVVLGDSNSIYPW 178

Db 172 DIVVLGDSNSIYPW 186

RESULT 68

AAB30928

ID AAB30928 standard; peptide; 15 AA.

XX

AC AAB30928;

XX

DT 02-APR-2001 (first entry)

XX

DE Antigenic peptide derived from a human alpha11 integrin chain.

XX

KW Human; integrin; alpha11 subunit; fibroblast; muscle cell; chondrocyte;  
 KW osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;  
 KW wound healing; trauma; rheumatoid arthritis; osteoarthritis;  
 KW osteoporosis; cartilage damage; bone damage; cartilage.

XX

OS Homo sapiens.

XX

FN WO200075187-A1.

XX

PD 14-DEC-2000.

XX

PP 31-MAY-2000; 2000WO-SE001135.

XX

PR 03-JUN-1999; 99SE-00002056.

XX

PA (ACT1-) ACTIVE BIOTECH AB.

XX

PI Gullberg D;

XX

DR WPI; 2001-071061/08.

XX

PT Integrin subunit alpha 11 or integrin heterodimer comprising subunit  
 PT alpha 11 in association with subunit beta, useful for treating muscle  
 PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.

XX

PS Disclosure; Page 17; 79pp; English.

XX

CC The present sequence is derived from the cytoplasmic domain of the human  
 CC integrin subunit, designated alpha11. The alpha11 polynucleotide and  
 CC polypeptide are useful as markers of cell target molecules, such as  
 CC fibroblasts, muscle cells, chondrocytes, osteoblasts, mesenchymally  
 CC derived cells or stem cells. They are also used for determining the  
 CC differential stage of cells during differentiation, development in  
 CC pathological conditions, in tissue regeneration, in transplantation or in  
 CC therapeutic and physiological repair of tissues. The pathological  
 CC conditions involving subunit alpha11 are selected from damage of cells,  
 CC muscle dystrophy, fibrosis, wound healing, trauma, rheumatoid arthritis,  
 CC osteoarthritis and osteoporosis, damage of cartilage and bone, and  
 CC cartilage and bone diseases. The polypeptide is useful for detecting the  
 CC formation of cartilage during embryonic development, for detecting  
 CC physiological therapeutic repair of cartilage and muscle, for selection  
 CC and analysis, or for sorting, isolating or purification of chondrocytes  
 CC and muscle cells, for detecting regeneration of cartilage or chondrocytes  
 CC during transplantation of cartilage or chondrocytes during

CC transplantation of cartilage or chondrocytes, respectively, or of muscle  
 CC or muscle cells during transplantation of muscle or muscle cells,  
 CC respectively, and for studies of differentiation or chondrocytes or  
 CC muscle cells

XX Sequence 15 AA;

Query Match 1.2%; Score 14; DB 4; Length 15;

Best Local Similarity 100.0%; Pred. No. 9.1e-06;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1175 REEPGLDPTPKVLE 1188

Db 2 REEPGLDPTPKVLE 15

RESULT 69

AAV07728

ID AAV07728 standard; protein; 1183 AA.

XX

AC AAV07728;

XX

DT 01-JUL-1999 (first entry)

XX

DE Armenian hamster alpha-1 integrin subunit protein.

XX

KW VEGF; tumour angiogenesis inhibition; vascular endothelial growth factor;  
 KW integrin cell surface receptor; capillary; blood vessel; hamster;  
 KW alpha-1 subunit; alpha-2 subunit.

XX

OS Cricetulus migratorius.

XX

PN WO9916465-A1.

XX

PD 08-APR-1999.

XX

PP 30-SEP-1997; 97WO-US017485.

XX

PR 30-SEP-1997; 97WO-US017485.

XX

PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX

PA (SENG/) SENG D R.

XX

PA (DETM/) DETMAR M.

XX

PA (CLAF/) CLAFFEY K P.

XX

PI Senger DR, Detmar M, Claffey KP;

XX

DR WPI; 1999-254930/21.

XX

PT Inhibition of tumor angiogenesis through interaction of vascular

XX

PS endotheial growth factor and integrin cell surface receptors.

XX

PS Disclosure; Fig 2A-C; 64pp; English.

XX

CC This invention describes a novel method for the inhibition of tumour  
 CC angiogenesis mediated by vascular endothelial growth factor (VEGF) and  
 CC integrin cell surface receptors expressed in vasculature of living  
 CC subjects. The method inhibits new capillary and new blood vessel  
 CC formation both within a tumour mass itself as well as in the immediately  
 CC adjacent blood vasculature surrounding the perimeter of the tumour mass.  
 CC Interaction and dependence upon VEGF to induce specific integrin  
 CC heterodimers in tumour angiogenesis provides a novel method for  
 CC inhibiting tumour angiogenesis, and unlike prior art relies on the  
 CC specific inter-relationship of VEGF and integrins, rather than  
 CC concentrating solely on one specific class of protein

XX

Sequence 1183 AA;

Query Match 1.2%; Score 14; DB 2; Length 1183;

Best Local Similarity 100.0%; Pred. No. 0.00042;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGDSNSIYP 177

```

Db      174 DIVLDGNSIYP 187
|||||
RESULT 70
AAU76862
ID AAU76862 standard; protein; 195 AA.
XX
AC AAU76862;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human integrin alpha subunit Alpha 10 variant A domain.
XX
KW Human; integrin alpha subunit; A domain; Alpha 10; integrin beta subunit;
KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
KW ischaemia-reperfusion injury; immune complex; parasitic disease; mutant;
KW antinflammatory; vasotropic; antiparasitic; vulnery; gene therapy;
KW muten.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 193
FT FT /note= "Wild-type ile substituted by any other amino
FT FT acid"
XX
PN WO200209737-A1.
XX
PD 07-FEB-2002.
XX
PF 31-JUL-2001; 2001WO-US023957.
XX
PR 31-JUL-2000; 2000US-0221950P.
PR 11-JAN-2001; 2001US-00758493.
PR 13-MAR-2001; 2001US-00805354.
XX
PA (GHEO ) GEN HOSPITAL CORP.
XX
PI Arnaout AM, Li R, Xiong J;
XX
DR WPI; 2002-188687/24.
XX
PT Novel high affinity integrin polypeptide useful for treating restenosis
PT and parasitic diseases, comprises all or part of variant integrin alpha
PT subunit A domain or variant integrin beta subunit A-like domain.
XX
PS Claim 53; Page; 55pp; English.
XX
CC The invention relates to a high affinity integrin polypeptide comprising
CC all or part of a variant integrin alpha subunit A domain or a variant
CC integrin beta subunit A-like domain. The polypeptide, preferably the
CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
CC V at residue 315 and A at residue 320 have been replaced by C, is useful
CC for determining if a test compound is a candidate compound for binding to
CC CD11b or for treating an inflammatory disorder, by contacting a test
CC compound with the polypeptide and determining if the test compound binds
CC to the polypeptide. The integrin subunits are useful for reducing
CC skeletal muscle injury, for treating disorders caused by ischaemia-
CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
CC to purify variant integrin polypeptide ligands and as bait proteins in
CC two-hybrid or three-hybrid assays. This sequence represents a human
CC integrin alpha subunit Alpha 10 variant A domain. Note: This variant
CC sequence is not featured in the specification but has been derived from
CC the wild-type protein shown in AAU76853
XX
SQ Sequence 195 AA;
XX
Query Match 1.1%; Score 13; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.00085;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      166 VIVLDGNSIYPW 178
|||||
RESULT 71
AAU76853
ID AAU76853 standard; protein; 195 AA.
XX
AC AAU76853;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human integrin alpha subunit Alpha 10 A domain.
XX
KW Human; integrin alpha subunit; A domain; Alpha 10; integrin beta subunit;
KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
KW ischaemia-reperfusion injury; immune complex; parasitic disease;
KW antinflammatory; vasotropic; antiparasitic; vulnery; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200209737-A1.
XX
PD 07-FEB-2002.
XX
PF 31-JUL-2001; 2001WO-US023957.
XX
PR 31-JUL-2000; 2000US-0221950P.
PR 11-JAN-2001; 2001US-00758493.
PR 13-MAR-2001; 2001US-00805354.
XX
PA (GHEO ) GEN HOSPITAL CORP.
XX
PI Arnaout AM, Li R, Xiong J;
XX
DR WPI; 2002-188687/24.
XX
PT Novel high affinity integrin polypeptide useful for treating restenosis
PT and parasitic diseases, comprises all or part of variant integrin alpha
PT subunit A domain or variant integrin beta subunit A-like domain.
XX
PS Example 2; Fig 5; 55pp; English.
XX
CC The invention relates to a high affinity integrin polypeptide comprising
CC all or part of a variant integrin alpha subunit A domain or a variant
CC integrin beta subunit A-like domain. The polypeptide, preferably the
CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
CC V at residue 315 and A at residue 320 have been replaced by C, is useful
CC for determining if a test compound is a candidate compound for binding to
CC CD11b or for treating an inflammatory disorder, by contacting a test
CC compound with the polypeptide and determining if the test compound binds
CC to the polypeptide. The integrin subunits are useful for reducing
CC skeletal muscle injury, for treating disorders caused by ischaemia-
CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
CC to purify variant integrin polypeptide ligands and as bait proteins in
CC two-hybrid or three-hybrid assays. This sequence represents the human
CC integrin alpha subunit Alpha 10 A domain
XX
SQ Sequence 195 AA;
XX
Query Match 1.1%; Score 13; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.00085;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      166 VIVLDGNSIYPW 178
|||||
RESULT 72
AAU76853
ID AAU76853 standard; protein; 195 AA.
XX
AC AAU76853;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human integrin alpha subunit Alpha 10 A domain.
XX
KW Human; integrin alpha subunit; A domain; Alpha 10; integrin beta subunit;
KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
KW ischaemia-reperfusion injury; immune complex; parasitic disease;
KW antinflammatory; vasotropic; antiparasitic; vulnery; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200209737-A1.
XX
PD 07-FEB-2002.
XX
PF 31-JUL-2001; 2001WO-US023957.
XX
PR 31-JUL-2000; 2000US-0221950P.
PR 11-JAN-2001; 2001US-00758493.
PR 13-MAR-2001; 2001US-00805354.
XX
PA (GHEO ) GEN HOSPITAL CORP.
XX
PI Arnaout AM, Li R, Xiong J;
XX
DR WPI; 2002-188687/24.
XX
PT Novel high affinity integrin polypeptide useful for treating restenosis
PT and parasitic diseases, comprises all or part of variant integrin alpha
PT subunit A domain or variant integrin beta subunit A-like domain.
XX
PS Example 2; Fig 5; 55pp; English.
XX
CC The invention relates to a high affinity integrin polypeptide comprising
CC all or part of a variant integrin alpha subunit A domain or a variant
CC integrin beta subunit A-like domain. The polypeptide, preferably the
CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
CC V at residue 315 and A at residue 320 have been replaced by C, is useful
CC for determining if a test compound is a candidate compound for binding to
CC CD11b or for treating an inflammatory disorder, by contacting a test
CC compound with the polypeptide and determining if the test compound binds
CC to the polypeptide. The integrin subunits are useful for reducing
CC skeletal muscle injury, for treating disorders caused by ischaemia-
CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
CC to purify variant integrin polypeptide ligands and as bait proteins in
CC two-hybrid or three-hybrid assays. This sequence represents the human
CC integrin alpha subunit Alpha 10 A domain
XX
SQ Sequence 195 AA;
XX
Query Match 1.1%; Score 13; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.00085;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

AAV32243	
ID	AAV32243 standard; protein; 1132 AA.
XX	
AC	
XX	
DT	15-FEB-2000 (first entry)
DE	Human integrin subunit alpha-10 splice variant.
XX	
KW	Integrin alpha-10; ISa10; human; trauma; rheumatoid arthritis;
KW	osteoarthritis; osteoarthritis; cancer; atherosclerosis; inflammation;
KW	therapy; cartilage; chondrocyte; osteoblast; fibroblast; vaccine; marker;
XW	splice variant.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Peptide
FT	Location/Qualifiers
FT	1..22
FT	/note= "signal peptide"
FT	Protein
FT	23..1132
FT	/note= "mature protein"
XX	
PX	WO9951639-A1.
PN	
PD	14-OCT-1999.
PD	
PF	31-MAR-1999; 99WO-SE000544.
PR	
PR	02-APR-1998; 98SE-00001164.
PR	28-JAN-1999; 99SE-00000319.
XX	
PA	(ACTI-) ACTIVE BIOTECH AB.
XX	
PI	Lundgren-Akerlund B;
DR	
DR	WFI; 2000-052639/04.
DR	N-PSDB; AAZ34720.
PT	New isolated integrin subunit alpha-10, used as a marker or target
PT	molecule for cells during development, regeneration and pathological
PT	conditions, e.g. arthritis, osteoarthritis, cancer, atherosclerosis or
PT	inflammation.
XX	
PS	Claim 1; Page 43-48; 90pp: English.
XX	
CC	This sequence represents a splice variant of novel human chondrocyte
CC	integrin subunit alpha-10 (ISa10). It is identical to ISa10 (see
CC	AAV32242) except for deletion of amino acids 975-986. The invention
CC	relates to a recombinant or isolated integrin heterodimer comprising the
CC	alpha10 subunit in association with subunit beta (especially beta-1). The
CC	heterodimer, subunit alpha-10 or splice variant can be used as a marker
CC	or target of all types of cells, e.g. of chondrocytes, osteoblasts and
CC	fibroblasts. They can also be used for treating pathological conditions
CC	involving ISa10, such as damage to cartilage, trauma, rheumatoid
CC	arthritis or osteoarthritis; for detecting the formation of cartilage
CC	during embryonal development, physiological or therapeutic repair of
CC	cartilage, or detecting regeneration of cartilage or chondrocytes during
CC	transplantation of cartilage or chondrocytes; for selection and analysis
CC	or for sorting, isolating or purification of chondrocytes and for in
CC	vivo studies of differentiation of chondrocytes; and as a target for
CC	anti-adhesive drugs or molecules in tendon, ligament, skeletal muscle or
CC	other tissues where adhesion impairs the function of the tissue (all
CC	claimed). ISa10 binding entities can be used to determine the
CC	differentiation-state of cells during embryonic development, angiogenesis
CC	or development of cancer, in pathological conditions such as rheumatoid
CC	arthritis, osteoarthritis or cancer, in tissue regeneration or in
CC	therapeutic and physiological repair of cartilage (claimed). A
CC	vaccine comprising the integrin heterodimer or subunit alpha-10 is also
CC	claimed. ISa10 polynucleotides, vectors, host cells and methods of
CC	producing recombinant ISa10 are also claimed
XX	
SQ	Sequence 1132 AA;

Query Match	1.1%; Score 13; DB 3; Length 1132;
Best Local Similarity	100.0%; Pred. No. 0.0039;
Matches	13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	166 VIVLDGSGNSIYPW 178
DB	169 VIVLDGSGNSIYPW 181
RESULT 73	
AAB64657	
ID	AAB64657 standard; protein; 1152 AA.
XX	
AC	AAB64657;
XX	
DT	22-MAR-2001 (first entry)
XX	
DE	Human secreted protein BLAST search protein SEQ ID NO: 167.
XX	
KW	Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW	anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;
KW	vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW	cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW	neurological disease; infection; human; secreted protein.
XX	
OS	Homo sapiens.
XX	
PN	WO200077197-A1.
XX	
PD	21-DEC-2000.
XX	
PF	01-JUN-2000; 2000WO-US014934.
PR	
PR	11-JUN-1999; 99US-0138599P.
XX	
PA	{HUMA-} HUMAN GENOME SCI INC.
PA	{ROSE/} ROSEN C.A.
XX	
PI	Rosen CA, Ruben SM, Komatsoulis GA;
XX	
DR	WPI; 2001-032312/04.
XX	
PT	Isolated nucleic acid molecule encoding a human secreted protein is used
PT	in preventing, treating or ameliorating a medical condition.
XX	
PS	Disclosure; Page 543-546; 558pp; English.
XX	
CC	The invention relates to the isolation of genes AAF32757-FJ2803 encoding
CC	the human secreted proteins AAB64549-B64594. The sequence is a search
CC	result from a BLASTX homology search. The genes and proteins are useful
CC	for preventing, ameliorating or treating medical conditions, e.g. by
CC	protein or gene therapy. The genes are isolated from a range of human
CC	tissues disclosed in the specification. The nucleic acids, proteins,
CC	antibodies and (antagonists are useful in the diagnosis, treatment and
CC	prevention of: (a) cancer, e.g. breast and ovarian cancer, and other
CC	cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal
CC	tract, liver, lung, or urogenital; (b) immune disorders, e.g. Addison's
CC	disease, allergies, autoimmune haemolytic anaemia, autoimmune
CC	thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis,
CC	rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders
CC	such as myocardial ischaemias; (d) wound healing; (e) neurological
CC	diseases, e.g. cerebral anoxia and epilepsy; and (f) infectious diseases
CC	such as viral, bacterial, fungal and parasitic infections
XX	
SQ	Sequence 1152 AA;

Query Match	1.1%; Score 13; DB 4; Length 1152;
Best Local Similarity	100.0%; Pred. No. 0.004;
Matches	13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	166 VIVLDGSGNSIYPW 178
DB	169 VIVLDGSGNSIYPW 181
Sequence 1132 AA;	

RESULT 74  
 AAB64658  
 ID AAB64658 standard; protein; 1152 AA.  
 AC AAB64658;  
 XX  
 DT 22-MAR-2001 (first entry)  
 XX  
 DE Human secreted protein BLAST search protein SEQ ID NO: 168.  
 XX  
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;  
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200077197-A1.  
 XX  
 PD 21-DEC-2000.  
 XX  
 PF 01-JUN-2000; 2000WO-US014934.  
 XX  
 PR 11-JUN-1999; 99US-0138599P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.  
 XX  
 PI Rosen CA, Ruben SM, Komatsculis GA;  
 XX  
 DR WPI; 2001-032312/04.  
 XX  
 PT Isolated nucleic acid molecule encoding a human secreted protein is used  
 PT in preventing, treating or ameliorating a medical condition.  
 XX  
 PS Disclosure; Page 547-551; 558pp; English.  
 XX  
 CC The invention relates to the isolation of genes AAB6458-P32803 encoding  
 CC the human secreted proteins AAB6459-B64594. The sequence is used as a  
 CC query sequence for doing BLASTX searches to identify homologous  
 CC sequences. The genes and proteins are useful for preventing, ameliorating  
 CC or treating medical conditions, e.g. by protein or gene therapy. The  
 CC genes are isolated from a range of human tissues disclosed in the  
 CC specification. The nucleic acids, proteins, antibodies and (ant)agonists  
 CC are useful in the diagnosis, treatment and prevention of: (a) cancer,  
 CC e.g. breast and ovarian cancer, and other cancers of the adrenal gland,  
 CC bone, bone marrow, breast, gastrointestinal tract, liver, lung, or  
 CC urogenital; (b) immune disorders e.g. Addison's disease, allergies,  
 CC autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,  
 CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d)  
 CC wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections  
 XX  
 SQ Sequence 1152 AA;  
 Query Match 1.1%; Score 13; DB 4; Length 1152;  
 Best Local Similarity 100.0%; Pred. No. 0.004;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 166 VIVLDGNSIYPW 178  
 DB 169 VIVLDGNSIYPW 181  
 |||||  
 RESULT 75  
 AAY32242  
 ID AAY32242 standard; protein; 1167 AA.  
 XX

AC AAY32242;  
 XX  
 DT 15-FEB-2000 (first entry)  
 XX  
 DE Human integrin subunit alpha-10.  
 XX  
 KW Integrin alpha-10; ISa10; human; trauma; rheumatoid arthritis;  
 KW osteoarthritis; osteoarthritis; cancer; atherosclerosis; inflammation;  
 KW therapy; cartilage; chondrocyte; osteoblast; fibroblast; vaccine; marker.  
 XX  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT Peptide 1..22  
 FT Protein /note= "signal peptide"  
 FT /note= "1145"  
 FT Domain /note= "mature protein"  
 FT /note= "23..1120"  
 FT /note= "extracellular domain"  
 FT /note= "98"  
 FT /note= "N-glycosylated"  
 FT Domain 162..359  
 FT /note= "I-domain"  
 FT Modified-site 336  
 FT /note= "N-glycosylated"  
 FT Modified-site 364  
 FT /note= "N-glycosylated"  
 FT Binding-site 494..502  
 FT /note= "cation binding site motif"  
 FT Binding-site 558..566  
 FT /note= "cation binding site motif"  
 FT Binding-site 620..628  
 FT /note= "cation binding site motif"  
 FT Modified-site 733  
 FT /note= "N-glycosylated"  
 FT Modified-site 839  
 FT /note= "N-glycosylated"  
 FT Modified-site 921  
 FT /note= "N-glycosylated"  
 FT Modified-site 1018  
 FT /note= "N-glycosylated"  
 FT Modified-site 1039  
 FT /note= "N-glycosylated"  
 FT Domain 1121..1145  
 FT /note= "transmembrane domain"  
 FT Domain 1122..1167  
 FT /note= "cytoplasmic domain, specifically claimed in Claim 21"  
 XX  
 PN WO9951639-A1.  
 XX  
 PD 14-OCT-1999.  
 XX  
 PP 31-MAR-1999; 99WO-SE000544.  
 XX  
 PR 02-APR-1998; 98SE-00001164.  
 PR 28-JAN-1999; 99SE-00000319.  
 XX  
 PA (ACTI-) ACTIVE BIOTECH AB.  
 XX  
 PI Lundgren-Akerlund E;  
 DR WPI; 2000-052639/04.  
 DR N-PSDB; AAZ34719.  
 XX  
 PT New isolated integrin subunit alpha-10, used as a marker or target  
 PT molecule for cells during development, regeneration and pathological  
 PT conditions, e.g. arthritis, osteoarthritis, cancer, atherosclerosis or  
 PT inflammation.  
 XX  
 PS Claim 1; Fig 6; 90pp; English.  
 XX  
 CC This sequence represents novel human chondrocyte integrin subunit alpha-

```

Query Match      1.1%  Score 13;  DB 3;  Length 1167;
Best Local Similarity 100.0%;  Pred. NO. 0.00041;
Matches 13;  Conservative 0;  Mismatches 0;  Gaps 0;

QY  166  VIVLDGNSIYPW 178
      |||||
Db   169  VIVLDGNSIYPW 181

```

Search completed: June 24, 2004, 18:03:55  
Job time : 72 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 18:05:07 ; Search time 53 Seconds  
(without alignments)  
6328.087 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 1188

Sequence: 1 MDLPRGLVVALSLWPGFT.....FRSARRRREPGLDTPPKVLE 1188

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1163542 seqs, 282313646 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications AA:  
1: /cgm2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgm2\_6/ptodata/2/pubpaa/RCT\_NEW\_PUB.pep.\*  
3: /cgm2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgm2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgm2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgm2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgm2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgm2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgm2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgm2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgm2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgm2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgm2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgm2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgm2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgm2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgm2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgm2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	926	77.9	1188	15	US-10-291-265-338
2	911	76.7	1188	15	US-10-291-265-810
3	911	76.7	1189	10	US-09-984-130-35
4	911	76.7	1189	10	US-09-836-353A-35
5	911	76.7	1189	12	US-10-262-839-4
6	810	68.2	1034	10	US-09-984-130-43
7	810	68.2	1034	10	US-09-836-353A-43
8	753	63.4	1120	12	US-10-262-839-6
9	489	41.2	707	9	US-09-764-870-313
10	489	41.2	707	14	US-10-145-540-313
11	245	16.2	437	15	US-10-108-260A-3386
12	193	16.2	193	10	US-09-805-354-8
13	193	16.2	193	12	US-09-758-493-8
14	193	16.2	193	14	US-10-144-259-8
15	88	7.4	103	9	US-09-764-870-472

16	88	7.4	103	11	US-09-764-875-1193	Sequence 1193, Ap
17	88	7.4	103	14	US-10-125-540-472	Sequence 472, App
18	42	3.5	42	14	US-10-144-259-25	Sequence 25, Appl
19	42	3.5	545	10	US-09-866-050A-500	Sequence 500, App
20	42	3.5	688	10	US-09-866-050A-624	Sequence 624, App
21	42	3.5	696	10	US-09-866-050A-501	Sequence 501, App
22	33	2.8	33	10	US-09-984-130-102	Sequence 102, App
23	33	2.8	33	10	US-09-836-353A-102	Sequence 102, App
24	21	1.8	158	10	US-09-866-050A-391	Sequence 391, App
25	17	1.4	85	10	US-09-866-050A-390	Sequence 390, App
26	15	1.3	148	9	US-09-764-870-284	Sequence 284, App
27	15	1.3	148	9	US-09-764-870-444	Sequence 444, App
28	15	1.3	148	14	US-10-125-540-284	Sequence 284, App
29	15	1.3	148	14	US-10-125-540-444	Sequence 444, App
30	15	1.3	192	12	US-10-346-863-57	Sequence 57, Appl
31	15	1.3	192	16	US-10-474-832-59	Sequence 59, Appl
32	15	1.3	192	16	US-10-474-832-60	Sequence 60, Appl
33	15	1.3	192	16	US-10-474-832-61	Sequence 61, Appl
34	15	1.3	195	10	US-09-805-354-5	Sequence 5, Appl
35	15	1.3	195	12	US-09-758-493-5	Sequence 5, Appl
36	15	1.3	195	14	US-10-144-259-5	Sequence 5, Appl
37	15	1.3	212	9	US-09-996-738-5	Sequence 5, Appl
38	15	1.3	212	9	US-09-996-738-6	Sequence 6, Appl
39	15	1.3	214	12	US-10-625-260-5	Sequence 5, Appl
40	15	1.3	214	12	US-10-625-260-6	Sequence 6, Appl
41	15	1.3	214	12	US-10-625-260-9	Sequence 9, Appl
42	15	1.3	214	13	US-10-061-658-5	Sequence 5, Appl
43	15	1.3	214	13	US-10-061-658-6	Sequence 6, Appl
44	15	1.3	214	13	US-10-061-658-9	Sequence 9, Appl
45	15	1.3	214	16	US-10-474-832-63	Sequence 63, Appl
46	15	1.3	214	16	US-10-474-832-64	Sequence 64, Appl
47	15	1.3	1151	10	US-09-984-130-103	Sequence 103, App
48	15	1.3	1151	10	US-09-836-353A-103	Sequence 103, App
49	15	1.3	1179	12	US-09-918-715-250	Sequence 250, App
50	15	1.3	1180	12	US-09-918-715-307	Sequence 307, App
51	13	1.1	195	10	US-09-805-354-7	Sequence 7, Appl
52	13	1.1	195	12	US-09-758-493-7	Sequence 7, Appl
53	13	1.1	195	14	US-10-144-259-7	Sequence 7, Appl
54	11	0.9	43	14	US-10-144-259-26	Sequence 26, Appl
55	11	0.9	1147	12	US-10-336-603A-42	Sequence 42, Appl
56	11	0.9	1181	12	US-10-211-462-187	Sequence 187, App
57	11	0.9	1181	14	US-10-160-354-2	Sequence 2, Appl
58	11	0.9	1181	15	US-10-295-027-1286	Sequence 1286, Ap
59	9	0.8	164	15	US-10-108-260A-3415	Sequence 3415, Ap
60	9	0.8	312	16	US-10-038-854-102	Sequence 102, App
61	9	0.8	371	15	US-10-369-493-2470	Sequence 2470, Ap
62	8	0.7	17	9	US-09-350-259-20	Sequence 20, Appl
63	8	0.7	17	10	US-09-891-943-20	Sequence 20, Appl
64	8	0.7	43	14	US-10-144-259-27	Sequence 27, Appl
65	8	0.7	67	9	US-09-978-295A-85	Sequence 85, Appl
66	8	0.7	67	9	US-09-978-697-85	Sequence 85, Appl
67	8	0.7	67	9	US-09-978-192A-85	Sequence 85, Appl
68	8	0.7	67	9	US-09-999-832A-85	Sequence 85, Appl
69	8	0.7	67	10	US-09-978-189-85	Sequence 85, Appl
70	8	0.7	67	10	US-09-978-608A-85	Sequence 85, Appl
71	8	0.7	67	10	US-09-978-585A-85	Sequence 85, Appl
72	8	0.7	67	10	US-09-978-191A-85	Sequence 85, Appl
73	8	0.7	67	10	US-09-978-403A-85	Sequence 85, Appl
74	8	0.7	67	10	US-09-978-564A-85	Sequence 85, Appl
75	8	0.7	67	10	US-09-999-831A-85	Sequence 85, Appl
76	8	0.7	67	10	US-09-981-915A-85	Sequence 85, Appl
77	8	0.7	67	10	US-09-978-824-85	Sequence 85, Appl
78	8	0.7	67	10	US-09-918-585A-85	Sequence 85, Appl
79	8	0.7	67	10	US-09-978-423A-85	Sequence 85, Appl
80	8	0.7	67	10	US-09-978-193A-85	Sequence 85, Appl
81	8	0.7	67	10	US-09-999-830A-85	Sequence 85, Appl
82	8	0.7	67	10	US-09-978-757A-85	Sequence 85, Appl
83	8	0.7	67	10	US-09-978-187B-85	Sequence 85, Appl
84	8	0.7	67	10	US-09-978-643A-85	Sequence 85, Appl
85	8	0.7	67	10	US-09-978-375A-85	Sequence 85, Appl
86	8	0.7	67	10	US-09-978-298A-85	Sequence 85, Appl
87	8	0.7	67	10	US-09-978-188A-85	Sequence 85, Appl
88	8	0.7	67	10	US-09-978-681A-85	Sequence 85, Appl

```
89 8 0.7 67 10 US-09-978-194A-85 Sequence 85, Appl
90 8 0.7 67 10 US-09-999-829A-85 Sequence 85, Appl
91 8 0.7 67 10 US-09-978-299A-85 Sequence 85, Appl
92 8 0.7 67 10 US-09-978-544A-85 Sequence 85, Appl
93 8 0.7 67 10 US-09-978-665A-85 Sequence 85, Appl
94 8 0.7 67 10 US-09-978-802A-85 Sequence 85, Appl
95 8 0.7 67 12 US-10-164-749A-85 Sequence 324, App
96 8 0.7 67 12 US-10-221-278-324 Sequence 85, Appl
97 8 0.7 67 12 US-09-999-831A-85 Sequence 85, Appl
98 8 0.7 67 12 US-10-013-917A-85 Sequence 85, Appl
99 8 0.7 67 12 US-09-999-834A-85 Sequence 85, Appl
100 8 0.7 67 12 US-10-162-521A-85 Sequence 85, Appl

ALIGNMENTS

RESULT 1
US-10-291-265-338
; Sequence 338, Application US/10291265
; Publication No. US20030232054A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides
; CURRENT APPLICATION NUMBER: US/10/291,265
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 338
; LENGTH: 1188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-265-338

Query Match 77.9%; Score 926; DB 15; Length 1188;
Best Local Similarity 99.88; Pred. No. 0;
Matches 1126; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDLPRLVAVWALS L P G F T D T F N M T R K P R V I P G S R T A P F G T V Q Q H D I S G N K W L V V G A 60
Db 1 MDLPRLVAVWALS L P G F T D T F N M T R K P R V I P G S R T A P F G T V Q Q H D I S G N K W L V V G A 60

QY 61 PLENTGYQKTDVYKCPVHGNCTKLNGLRVTLNYSVERKDNMRGLGLSLATNPKNPLA 120
Db 61 PLENTGYQKTDVYKCPVHGNCTKLNGLRVTLNYSVERKDNMRGLGLSLATNPKNPLA 120

QY 121 CSPLSWHECGSSYTTGMC SRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNIYPWVE 180
Db 121 CSPLSWHECGSSYTTGMC SRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSNIYPWVE 180

QY 181 VQHFNLNLKFFYIGPQIQGVVQYGEDVVFHFHNDVSVKDVVVEAASHIBQRCGTET 240
Db 181 VQHFNLNLKFFYIGPQIQGVVQYGEDVVFHFHNDVSVKDVVVEAASHIBQRCGTET 240

QY 241 RTAFGIEFARSEAFQGRKGAKKMIIVI TDGSHSDSPDLKVIQSERDNDVTRYAVVL 300
Db 241 RTAFGIEFARSEAFQGRKGAKKMIIVI TDGSHSDSPDLKVIQSERDNDVTRYAVVL 300

QY 301 GYNNRGINPETELNEIKYIASDPDDKHFNVTDEAALKDIDVALGDRIFSLGTTNKET 360
Db 301 GYNNRGINPETELNEIKYIASDPDDKHFNVTDEAALKDIDVALGDRIFSLGTTNKET 360

QY 361 SFGLEMSQTGFSSHVEDGVLLGAVGAYDNGAVLTKETSAGKVIPLRESYLKEFPEELKN 420
```

```
Db 361 SFGLEMSQTGFSSHVEDGVLLGAVGAYDNGAVLTKETSAGKVIPLRESYLKEFPEELKN 420
QY 421 HGAYLGTYTTSVVSQSGRQVYVAGAPRNFHTGKVLFTMHNNRSTI HQAMRGQOIGSYF 480
Db 421 HGAYLGTYTTSVVSQSGRQVYVAGAPRNFHTGKVLFTMHNNRSTI HQAMRGQOIGSYF 480
QY 481 GSEITSVDIDGSGVTDVLLVGAPMYFNEGRGRGKVVYVELRQNRVYVNGTLDKSHSYQVA 540
Db 481 GSEITSVDIDGSGVTDVLLVGAPMYFNEGRGRGKVVYVELRQNRVYVNGTLDKSHSYQVA 540
QY 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFGFRGSILKTIKQRTASELATG 600
Db 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFGFRGSILKTIKQRTASELATG 600
QY 601 LQYFCSSIHGQLDNLNEDGLIDLAVGALGNVILSRPVVQINASLHFPSKINI FHRDCK 660
Db 601 LQYFCSSIHGQLDNLNEDGLIDLAVGALGNVILSRPVVQINASLHFPSKINI FHRDCK 660
QY 661 RSGRDATCLAFICPTPIFLAPHQTTTGVIRYNATQDERRYT PRAHLDGGDRTNRAV 720
Db 661 RSGRDATCLAFICPTPIFLAPHQTTTGVIRYNATQDERRYT PRAHLDGGDRTNRAV 720
QY 721 LLSQSQBLCECERINFHVLTADYVKVTPSVYSLEDDPHGPMDDGWP TILRVSVYFVWG 780
Db 721 LLSQSQBLCECERINFHVLTADYVKVTPSVYSLEDDPHGPMDDGWP TILRVSVYFVWG 780
QY 781 CNEDEHCVDPDVLVDARSOLPTAMEYCORVLRKPAQDCSAYT LSFDTTVFIIESTQRVAV 840
Db 781 CNEDEHCVDPDVLVDARSOLPTAMEYCORVLRKPAQDCSAYT LSFDTTVFIIESTQRVAV 840
QY 841 EATLENRGENAYSTVLNISQSANLOFASLLOKEDSDGSEICVNEERLQKQVCNYSYPP 900
Db 841 EATLENRGENAYSTVLNISQSANLOFASLLOKEDSDGSEICVNEERLQKQVCNYSYPP 900
QY 901 RAKAKVAFRLDSEFSKSIPLHHLHLELAAGSDSNRSDTKEDNVAPLPHLKYEADVLFT 960
Db 901 RAKAKVAFRLDSEFSKSIPLHHLHLELAAGSDSNRSDTKEDNVAPLPHLKYEADVLFT 960
QY 961 RSSLSHYEVKLNSSLERYDGLGPPSCIFRIONLGLPFIHGMWKITIP IATRSNRL 1020
Db 961 RSSLSHYEVKLNSSLERYDGLGPPSCIFRIONLGLPFIHGMWKITIP IATRSNRL 1020
QY 1021 KLRFELTDEANTSCNMGNSTEYRPTVPEEDLRRAPQLNHSNDSVVS INCMIRLVPNQE 1080
Db 1021 KLRFELTDEANTSCNMGNSTEYRPTVPEEDLRRAPQLNHSNDSVVS INCMIRLVPNQE 1080
QY 1081 NFHLGNLWLSLKALKYKSMKIMVNAALQRFHSPPIFREDDPSROI 1128
Db 1081 NFHLGNLWLSLKALKYKSMKIMVNAALQRFHSPPIFREDDPSROI 1128
```

## RESULT 2

```
US-10-291-265-810
; Sequence 810, Application US/10291265
; Publication No. US20030232054A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang et al
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-017 (785)
; CURRENT APPLICATION NUMBER: US/10/291,265
; CURRENT FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
```



; SEQ ID NO 810  
; LENGTH: 1188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-291-265-810

Query Match 76.7%; Score 911; DB 15; Length 1188;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MDLPRGLVAVWALSINPFTDFNNMDTRKPRVPGSTAFPGYTVQOHDISGNKWLAVGA 60
DB 1 MDLPRGLVAVWALSINPFTDFNNMDTRKPRVPGSTAFPGYTVQOHDISGNKWLAVGA 60
QY 61 PLETHGYOKTGVYKCPVHGNCTKLNLRVTLSNVSRKDNMRGLSLATNPKNDSFLA 120
DB 61 PLETHGYOKTGVYKCPVHGNCTKLNLRVTLSNVSRKDNMRGLSLATNPKNDSFLA 120
QY 121 CSPWASHCEGSSYYTTCMCSRNSNFRSKTVAPALQRCQTYMDIVIVLDCSNSIYPWVE 180
DB 121 CSPWASHCEGSSYYTTCMCSRNSNFRSKTVAPALQRCQTYMDIVIVLDCSNSIYPWVE 180
QY 181 VQHEFLINILKKFYIGPGQIQGVVQYGEDVHVEHFLNDYRSVKDVVEAASHIEBORGTTET 240
DB 181 VQHEFLINILKKFYIGPGQIQGVVQYGEDVHVEHFLNDYRSVKDVVEAASHIEBORGTTET 240
QY 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGHSDSPDLKVIQOSRDNVTRYAVAVL 300
DB 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGHSDSPDLKVIQOSRDNVTRYAVAVL 300
QY 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVNTDEAALKDIVDALGDRIFSLGNTKNKET 360
DB 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVNTDEAALKDIVDALGDRIFSLGNTKNKET 360
QY 361 SPGLEMSQTGSSHHVVEGVLLGAVGDMNGAVLKETSAGKVIPLRSYLYKEPPEELKN 420
DB 361 SPGLEMSQTGSSHHVVEGVLLGAVGDMNGAVLKETSAGKVIPLRSYLYKEPPEELKN 420
QY 421 HGAYLGTYVTSVSSRQGRVYVAGAPRNHTGKVLFTMNNRSLTIHQMRGQOIGSYF 480
DB 421 HGAYLGTYVTSVSSRQGRVYVAGAPRNHTGKVLFTMNNRSLTIHQMRGQOIGSYF 480
QY 481 RFGSSIASVRDLNQDSYNDVVGAPLENDHAGAIYIFHGFRGSILKTPKQRIITASELATG 600
DB 481 RFGSSIASVRDLNQDSYNDVVGAPLENDHAGAIYIFHGFRGSILKTPKQRIITASELATG 600
QY 601 IQYFCCSIHQGLDNLNEDGLIDLAVGALGNVILMSRPVVQINASHLPFPPSKINIFHRDCK 660
DB 601 IQYFCCSIHQGLDNLNEDGLIDLAVGALGNVILMSRPVVQINASHLPFPPSKINIFHRDCK 660
QY 661 RSGRATCLAAFLCETPIFLASHFOITTVGIRYNATMDERYTPRAHLDEGGDRFTNRAV 720
DB 661 RSGRATCLAAFLCETPIFLASHFOITTVGIRYNATMDERYTPRAHLDEGGDRFTNRAV 720
QY 721 LLSQSQELCERINFLVLTADYVVKPVTFSVYSLEDPDHGPMLDDGWPTTLIRVSPVFWNG 780
DB 721 LLSQSQELCERINFLVLTADYVVKPVTFSVYSLEDPDHGPMLDDGWPTTLIRVSPVFWNG 780
QY 781 CNEDEHCVDPDLVDARSDLPTAMEYCORVLRKPAQDCSAYTSLPDTTVPFIITESTRQYAV 840
DB 781 CNEDEHCVDPDLVDARSDLPTAMEYCORVLRKPAQDCSAYTSLPDTTVPFIITESTRQYAV 840
QY 841 EATLENRGENAYSTVLNLSQSANLQFASLIQKEDSDGSIETCNBERLQKQCNVSPFF 900
DB 841 EATLENRGENAYSTVLNLSQSANLQFASLIQKEDSDGSIETCNBERLQKQCNVSPFF 900
QY 901 RAKAKVAFRLD 911
DB 901 RAKAKVAFRLD 911
```

## RESULT 3

US-09-984-130-35  
; Sequence 35, Application US/09984130  
; Publication No. US20030055231A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 12 Human Secreted Proteins  
; FILE REFERENCE: PF489P2  
; CURRENT APPLICATION NUMBER: US/09/984.130  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/243,792  
; PRIOR FILING DATE: 2000-10-30  
; PRIOR APPLICATION NUMBER: 09/836,353  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/198,407  
; PRIOR FILING DATE: 2000-04-19  
; PRIOR APPLICATION NUMBER: PCT/US99/25031  
; PRIOR FILING DATE: 1999-10-27  
; PRIOR APPLICATION NUMBER: 60/105,971  
; PRIOR FILING DATE: 1998-10-28  
; NUMBER OF SEQ ID NOS: 149  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 35  
; LENGTH: 1189  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-984-130-35

Query Match 76.7%; Score 911; DB 10; Length 1189;

Best Local Similarity 100.0%; Pred. No. 0;

```
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDLPRGLVAVWALSINPFTDFNNMDTRKPRVPGSTAFPGYTVQOHDISGNKWLAVGA 60
DB 1 MDLPRGLVAVWALSINPFTDFNNMDTRKPRVPGSTAFPGYTVQOHDISGNKWLAVGA 60
QY 61 PLETHGYOKTGVYKCPVHGNCTKLNLRVTLSNVSRKDNMRGLSLATNPKNDSFLA 120
DB 61 PLETHGYOKTGVYKCPVHGNCTKLNLRVTLSNVSRKDNMRGLSLATNPKNDSFLA 120
QY 121 CSPWASHCEGSSYYTTCMCSRNSNFRSKTVAPALQRCQTYMDIVIVLDCSNSIYPWVE 180
DB 121 CSPWASHCEGSSYYTTCMCSRNSNFRSKTVAPALQRCQTYMDIVIVLDCSNSIYPWVE 180
QY 181 VQHEFLINILKKFYIGPGQIQGVVQYGEDVHVEHFLNDYRSVKDVVEAASHIEBORGTTET 240
DB 181 VQHEFLINILKKFYIGPGQIQGVVQYGEDVHVEHFLNDYRSVKDVVEAASHIEBORGTTET 240
QY 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGHSDSPDLKVIQOSRDNVTRYAVAVL 300
DB 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGHSDSPDLKVIQOSRDNVTRYAVAVL 300
QY 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVNTDEAALKDIVDALGDRIFSLGNTKNKET 360
DB 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVNTDEAALKDIVDALGDRIFSLGNTKNKET 360
QY 361 SPGLEMSQTGSSHHVVEGVLLGAVGDMNGAVLKETSAGKVIPLRSYLYKEPPEELKN 420
DB 361 SPGLEMSQTGSSHHVVEGVLLGAVGDMNGAVLKETSAGKVIPLRSYLYKEPPEELKN 420
QY 421 HGAYLGTYVTSVSSRQGRVYVAGAPRNHTGKVLFTMNNRSLTIHQMRGQOIGSYF 480
DB 421 HGAYLGTYVTSVSSRQGRVYVAGAPRNHTGKVLFTMNNRSLTIHQMRGQOIGSYF 480
QY 481 GSEITTSVDIDGQDGVTDVLLVGA PMYFNEGRGKVVYVYELRQNRFFVNGTLKDSHSYQNA 540
DB 481 GSEITTSVDIDGQDGVTDVLLVGA PMYFNEGRGKVVYVYELRQNRFFVNGTLKDSHSYQNA 540
QY 541 RFGSSIASVRDLNQDSYNDVVGAPLENDHAGAIYIFHGFRGSILKTPKQRIITASELATG 600
DB 541 RFGSSIASVRDLNQDSYNDVVGAPLENDHAGAIYIFHGFRGSILKTPKQRIITASELATG 600
```

QY 601 LQYFGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQVQINASLHFEPSKINIPHRDCK 660  
DB 601 LQYFGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQVQINASLHFEPSKINIPHRDCK 660  
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDEGGDRFTNRVAV 720  
DB 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDEGGDRFTNRVAV 720  
QY 721 LLSGGQELCERINFHVLDTADYVKPVTFVSVEYSLEDDHGMDDGMPPTLRSVYPPFNG 780  
DB 721 LLSGGQELCERINFHVLDTADYVKPVTFVSVEYSLEDDHGMDDGMPPTLRSVYPPFNG 780  
QY 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTILSFDTTVEIIESTRQAVAV 840  
DB 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTILSFDTTVEIIESTRQAVAV 840  
QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECTVNEERLQKQVCNVSYPFF 900  
DB 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECTVNEERLQKQVCNVSYPFF 900  
QY 901 RAKAKVAFRLD 911  
DB 901 RAKAKVAFRLD 911

## RESULT 4

US-09-836-353A-35  
; Sequence 35, Application US/09836353A  
; Publication No. US20030129685A1

; GENERAL INFORMATION:

; APPLICANT: Ni et al.

; TITLE OF INVENTION: 12 Human Secreted Proteins

; FILE REFERENCE: PF489P1

; CURRENT APPLICATION NUMBER: US/09/836,353A

; CURRENT FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/198,407

; PRIOR FILING DATE: 2000-04-19

; PRIOR APPLICATION NUMBER: PCT/US99/25031

; PRIOR FILING DATE: 1999-10-27

; PRIOR FILING DATE: 1998-10-28

; NUMBER OF SEQ ID NOS: 147

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 35

; LENGTH: 1189

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-836-353A-35

Query Match 76.7%; Score 911; DB 10; Length 1189;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLPRGLVAVNALSLMPGFTDFTNMOTRKPRIIPGSRTPAFFGYTVOQHDISGNKMLVGA 60  
DB 1 MDLPRGLVAVNALSLMPGFTDFTNMOTRKPRIIPGSRTPAFFGYTVOQHDISGNKMLVGA 60  
QY 61 PLETNGYOKTGDVYKCPVTHGNCCKLNLGRVTLNVSERKDNRLGLSLATNPKNPSPLA 120  
DB 61 PLETNGYOKTGDVYKCPVTHGNCCKLNLGRVTLNVSERKDNRLGLSLATNPKNPSPLA 120  
QY 121 CSPLWSEHCSSYTTTGMCSSRVNSNFRPKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180  
DB 121 CSPLWSEHCSSYTTTGMCSSRVNSNFRPKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180  
QY 181 VOHFLINILKKEVIGRGQITGVGVQVGVQVGVQVGVQVGVQVGVQVGVQVGVQVGVQV 240  
DB 181 VOHFLINILKKEVIGRGQITGVGVQVGVQVGVQVGVQVGVQVGVQVGVQVGVQVGVQV 240  
QY 241 RTAFGEFARSFAFGKGRGKAKVMIVITDGSSESDSPLEKVIQOSERDNVTRVAVAVL 300  
DB 241 RTAFGEFARSFAFGKGRGKAKVMIVITDGSSESDSPLEKVIQOSERDNVTRVAVAVL 300

QY 301 GYNERGINPTEFLNEIKYIASDDDDKHFFNVNTDEAALKDIDVALGDRIIFSLEGTKNKET 360  
DB 301 GYNERGINPTEFLNEIKYIASDDDDKHFFNVNTDEAALKDIDVALGDRIIFSLEGTKNKET 360  
QY 361 SFGLEMSQTGSSSHVDEGVLLGAVGAYDMNGAVLKET'SAGKVIPLRSYILKEPPEELKN 420  
DB 361 SFGLEMSQTGSSSHVDEGVLLGAVGAYDMNGAVLKET'SAGKVIPLRSYILKEPPEELKN 420  
QY 421 HGAYLGYVTWTSVWASRQGVVAVAGAPRNFHTGKVLPTMHNHRSITTHQARGGQOIGSYF 480  
DB 421 HGAYLGYVTWTSVWASRQGVVAVAGAPRNFHTGKVLPTMHNHRSITTHQARGGQOIGSYF 480  
QY 481 GSEITSVDIDGQVTDVLLVAGAPMYFNEGRERKGVYVYVELRQNRFPVYNGTLKDSHSYQNA 540  
DB 481 GSEITSVDIDGQVTDVLLVAGAPMYFNEGRERKGVYVYVELRQNRFPVYNGTLKDSHSYQNA 540  
QY 541 RFGSSIASVRDLNODSYNDVYVVGAPLENDHAGAIYIPIHGFSGSILKTPKQITASELATG 600  
DB 541 RFGSSIASVRDLNODSYNDVYVVGAPLENDHAGAIYIPIHGFSGSILKTPKQITASELATG 600  
QY 601 LQYFGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQVQINASLHFEPSKINIPHRDCK 660  
DB 601 LQYFGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQVQINASLHFEPSKINIPHRDCK 660  
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDEGGDRFTNRVAV 720  
DB 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDEGGDRFTNRVAV 720  
QY 721 LLSGGQELCERINFHVLDTADYVKPVTFVSVEYSLEDDHGMDDGMPPTLRSVYPPFNG 780  
DB 721 LLSGGQELCERINFHVLDTADYVKPVTFVSVEYSLEDDHGMDDGMPPTLRSVYPPFNG 780  
QY 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTILSFDTTVEIIESTRQAVAV 840  
DB 781 CNEDEHCVDPDLVLDARSDLPTAMEYCORVLRKPAQDCSAYTILSFDTTVEIIESTRQAVAV 840  
QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECTVNEERLQKQVCNVSYPFF 900  
DB 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECTVNEERLQKQVCNVSYPFF 900  
QY 901 RAKAKVAFRLD 911  
DB 901 RAKAKVAFRLD 911

## RESULT 5

US-10-262-839-4

; Sequence 4, Application US/10262839

; Publication No. US20040038877A1

; GENERAL INFORMATION:

; APPLICANT: Alsbrook, John,

; APPLICANT: Anderson, David W.,

; APPLICANT: Boldog, Ferenc,

; APPLICANT: Burgess, Catherine,

; APPLICANT: Catterton, Elina,

; APPLICANT: Edinger, Shlomit,

; APPLICANT: Ellerman, Karen,

; APPLICANT: Gerlach, Valerie,

; APPLICANT: Gorman, Linda,

; APPLICANT: Guo, Xiaojia,

; APPLICANT: Ji, Weizhen,

; APPLICANT: Kekuda, Ramesh,

; APPLICANT: Leach, Martin,

; APPLICANT: Li, Li,

; APPLICANT: Miller, Charles,

; APPLICANT: Patturajan, Meera,

; APPLICANT: Reiger, Daniel,

; APPLICANT: Rothenberg, Mark,

; APPLICANT: Shinkets, Richard,

; APPLICANT: Smithson, Glennda,

; APPLICANT: Spyttek, Kimberly,

; APPLICANT: Taupier, Raymond, Jr.,

```

; APPLICANT: Vernet, Corine,
; APPLICANT: Voss, Edward,
; APPLICANT: Zethusen, Brian,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-462A
; CURRENT APPLICATION NUMBER: US/10/262,839
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,101
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/371,972
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/327,342
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/328,044
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/374,738
; PRIOR FILING DATE: 2002-04-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 4
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-839-4

```

```

Query Match      76.7%; Score 911; DB 12; Length 1189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLPRLVAVALSALPQGTDTFNDTRKPRVPGSRPTAFPGYTVQOHDISGNKWLAVGA 60
DB 1 MDLPRLVAVALSALPQGTDTFNDTRKPRVPGSRPTAFPGYTVQOHDISGNKWLAVGA 60
QY 61 PLETNGYQKTGDVYKCPVHGNCTKLNIGRVTLNVSERKNNRGLSLATNPKNNSFLA 120
DB 61 PLETNGYQKTGDVYKCPVHGNCTKLNIGRVTLNVSERKNNRGLSLATNPKNNSFLA 120
QY 121 CSPLMSHECGSSYYTGMCNRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
DB 121 CSPLMSHECGSSYYTGMCNRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
QY 181 VOHFLNLTILKKEVIGPGQIOGVWQYQGVDDVHEPHLNDYRSKQVWEAASHIEQGGTET 240
DB 181 VOHFLNLTILKKEVIGPGQIOGVWQYQGVDDVHEPHLNDYRSKQVWEAASHIEQGGTET 240
QY 241 RTAFGIEFARSEAFQKGRKGAKKMWIITDGSCHSDSPDLKVIQOQSERDNVTRYAVAVL 300
DB 241 RTAFGIEFARSEAFQKGRKGAKKMWIITDGSCHSDSPDLKVIQOQSERDNVTRYAVAVL 300
QY 301 GYNRRGINPETFLNEIKYIASDDDDKGFNNVTDEAALKDVIDALGDRIFLEGNTNKNET 360
DB 301 GYNRRGINPETFLNEIKYIASDDDDKGFNNVTDEAALKDVIDALGDRIFLEGNTNKNET 360
QY 361 SPGLEMSQTGSSHHVDEGVLLGAVGYDNGAVLKETSAKVTPLRESYLKEPPEELKN 420
DB 361 SPGLEMSQTGSSHHVDEGVLLGAVGYDNGAVLKETSAKVTPLRESYLKEPPEELKN 420
QY 421 HGAYLGTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMNNRSLTIHQAVRGOQOIGSYF 480
DB 421 HGAYLGTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMNNRSLTIHQAVRGOQOIGSYF 480

```

RESULT 6

```

US-09-984-130-43
; Sequence 43, Application US/09984130
; Publication No. US2003005231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: P2489P2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 43
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-43

```

```

Query Match      68.2%; Score 810; DB 10; Length 1034;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDLPRLVAVALSALPQGTDTFNDTRKPRVPGSRPTAFPGYTVQOHDISGNKWLAVGA 60
DB 1 MDLPRLVAVALSALPQGTDTFNDTRKPRVPGSRPTAFPGYTVQOHDISGNKWLAVGA 60
QY 61 PLETNGYQKTGDVYKCPVHGNCTKLNIGRVTLNVSERKNNRGLSLATNPKNNSFLA 120
DB 61 PLETNGYQKTGDVYKCPVHGNCTKLNIGRVTLNVSERKNNRGLSLATNPKNNSFLA 120

```

QY 121 CSPWASHCEGSSYYTTCMCVRNSNFRFSKTVAPALQRCOTYMDIVIVLDGNSNIYPWVE 180  
 DB 121 CSPWASHCEGSSYYTTCMCVRNSNFRFSKTVAPALQRCOTYMDIVIVLDGNSNIYPWVE 180  
 QY 181 VQHFLINILKFKFYIGPQIQGVVQGVGDDVHFEHLNDYRSVKDVVRAASHIEQGGTET 240  
 DB 181 VQHFLINILKFKFYIGPQIQGVVQGVGDDVHFEHLNDYRSVKDVVRAASHIEQGGTET 240  
 QY 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSHPDPLEKVIQOQSERDNVTRYAVAVL 300  
 DB 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSHPDPLEKVIQOQSERDNVTRYAVAVL 300  
 QY 301 GYNRRGINPETPNEIKYIASDDPKHFNVTDEAALKDIYDALGDRIFSLSEGTNNKET 360  
 DB 301 GYNRRGINPETPNEIKYIASDDPKHFNVTDEAALKDIYDALGDRIFSLSEGTNNKET 360  
 QY 361 SFGLMSQTFSSHHVEDGVLLGAVGAYDMNGAVLKEYSAGKVIPLRESYLKEPPELKN 420  
 DB 361 SFGLMSQTFSSHHVEDGVLLGAVGAYDMNGAVLKEYSAGKVIPLRESYLKEPPELKN 420  
 QY 421 HGAYLGYYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMNNRSLTIHQMRGQOIGSYF 480  
 DB 421 HGAYLGYYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMNNRSLTIHQMRGQOIGSYF 480  
 QY 481 GSEITSDIDGQVTDVLLVGAAPMYFNEGRERKGVYVELRQNRVYNGTLKDSHSYQNA 540  
 DB 481 GSEITSDIDGQVTDVLLVGAAPMYFNEGRERKGVYVELRQNRVYNGTLKDSHSYQNA 540  
 QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHEAGAIYIFHGFSGSILKTPKORITASELATG 600  
 DB 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHEAGAIYIFHGFSGSILKTPKORITASELATG 600  
 QY 601 LQYFGCSIHQGLDNLNEDGLDLAVGALGNVILWSRPVQINASLHFEPSKINIPIHEDCK 660  
 DB 601 LQYFGCSIHQGLDNLNEDGLDLAVGALGNVILWSRPVQINASLHFEPSKINIPIHEDCK 660  
 QY 721 LLSSGQELCERINFHVLDTADYVKPTFSVEYSLEDDPHGPMDDGWTTLRVSVPFWNG 780  
 DB 721 LLSSGQELCERINFHVLDTADYVKPTFSVEYSLEDDPHGPMDDGWTTLRVSVPFWNG 780  
 QY 781 CNEDEHCVDPDLVLDARSDLPTAMEYQORVLRKPAQDCSAYTILSFDTTFFIIESTRQVAV 840  
 DB 781 CNEDEHCVDPDLVLDARSDLPTAMEYQORVLRKPAQDCSAYTILSFDTTFFIIESTRQVAV 840  
 QY 841 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGSECTVNEERLQKQVNCVSYPPF 900  
 DB 841 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGSECTVNEERLQKQVNCVSYPPF 900  
 QY 901 RAKAKVAFRLD 911  
 DB 901 RAKAKVAFRLD 911

RESULT 7  
 US-09-836-353A-43  
 ; Sequence 43, Application US/09836353A  
 ; Publication No US20030129685A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NI et al.  
 ; TITLE OF INVENTION: 12 Human Secreted Proteins  
 ; FILE REFERENCE: PF499PI  
 ; CURRENT APPLICATION NUMBER: US/09/836,353A  
 ; CURRENT FILING DATE: 2001-04-18  
 ; PRIOR APPLICATION NUMBER: 60/198,407  
 ; PRIOR FILING DATE: 2000-04-19  
 ; PRIOR APPLICATION NUMBER: PCT/US99/25031  
 ; PRIOR FILING DATE: 1999-10-27  
 ; PRIOR APPLICATION NUMBER: 60/105,971  
 ; PRIOR FILING DATE: 1998-10-28

; NUMBER OF SEQ ID NOS: 147  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 43  
 ; LENGTH: 1034  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-836-353A-43

Query Match 68.2%; Score 810; DB 10; Length 1034;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDLPRLGVAVWALSILWPGFTDTENMDTKPRVIPCSTAPRGYVVOQHDISGNKMLVVG 60  
 DB 1 MDLPRLGVAVWALSILWPGFTDTENMDTKPRVIPCSTAPRGYVVOQHDISGNKMLVVG 60  
 QY 61 PLETTYGYQKTDVYKCPVIHGNCTKLMLGRVTLNSVSEKDNMRGLGSLATNPKDNSFLA 120  
 DB 61 PLETTYGYQKTDVYKCPVIHGNCTKLMLGRVTLNSVSEKDNMRGLGSLATNPKDNSFLA 120  
 QY 121 CSPWASHCEGSSYYTTCMCVRNSNFRFSKTVAPALQRCOTYMDIVIVLDGNSNIYPWVE 180  
 DB 121 CSPWASHCEGSSYYTTCMCVRNSNFRFSKTVAPALQRCOTYMDIVIVLDGNSNIYPWVE 180  
 QY 181 VQHFLINILKFKFYIGPQIQGVVQGVGDDVHFEHLNDYRSVKDVVRAASHIEQGGTET 240  
 DB 181 VQHFLINILKFKFYIGPQIQGVVQGVGDDVHFEHLNDYRSVKDVVRAASHIEQGGTET 240  
 QY 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSHPDPLEKVIQOQSERDNVTRYAVAVL 300  
 DB 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGSHPDPLEKVIQOQSERDNVTRYAVAVL 300  
 QY 301 GYNRRGINPETPNEIKYIASDDPKHFNVTDEAALKDIYDALGDRIFSLSEGTNNKET 360  
 DB 301 GYNRRGINPETPNEIKYIASDDPKHFNVTDEAALKDIYDALGDRIFSLSEGTNNKET 360  
 QY 361 SFGLMSQTFSSHHVEDGVLLGAVGAYDMNGAVLKEYSAGKVIPLRESYLKEPPELKN 420  
 DB 361 SFGLMSQTFSSHHVEDGVLLGAVGAYDMNGAVLKEYSAGKVIPLRESYLKEPPELKN 420  
 QY 421 HGAYLGYYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMNNRSLTIHQMRGQOIGSYF 480  
 DB 421 HGAYLGYYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMNNRSLTIHQMRGQOIGSYF 480  
 QY 481 GSEITSDIDGQVTDVLLVGAAPMYFNEGRERKGVYVELRQNRVYNGTLKDSHSYQNA 540  
 DB 481 GSEITSDIDGQVTDVLLVGAAPMYFNEGRERKGVYVELRQNRVYNGTLKDSHSYQNA 540  
 QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHEAGAIYIFHGFSGSILKTPKORITASELATG 600  
 DB 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHEAGAIYIFHGFSGSILKTPKORITASELATG 600  
 QY 601 LQYFGCSIHQGLDNLNEDGLDLAVGALGNVILWSRPVQINASLHFEPSKINIPIHEDCK 660  
 DB 601 LQYFGCSIHQGLDNLNEDGLDLAVGALGNVILWSRPVQINASLHFEPSKINIPIHEDCK 660  
 QY 661 RSGRDATCLAAFLCFTPIFLAPHQTTVGIRYNATMDERRYTPRAHLDEGGDRFTNRAV 720  
 DB 661 RSGRDATCLAAFLCFTPIFLAPHQTTVGIRYNATMDERRYTPRAHLDEGGDRFTNRAV 720  
 QY 721 LLSSGQELCERINFHVLDTADYVKPTFSVEYSLEDDPHGPMDDGWTTLRVSVPFWNG 780  
 DB 721 LLSSGQELCERINFHVLDTADYVKPTFSVEYSLEDDPHGPMDDGWTTLRVSVPFWNG 780  
 QY 781 CNEDEHCVDPDLVLDARSDLPTAMEYQORVLRKPAQDCSAYTILSFDTTFFIIESTRQVAV 840  
 DB 781 CNEDEHCVDPDLVLDARSDLPTAMEYQORVLRKPAQDCSAYTILSFDTTFFIIESTRQVAV 840  
 QY 841 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGSECTVNEERLQKQVNCVSYPPF 900  
 DB 841 EATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGSECTVNEERLQKQVNCVSYPPF 900  
 QY 901 RAKAKVAFRLD 911

Db	90	CQTYMDIVILDGNSNIYPPVEVQHFPLINLKKEFYIGPQIQVGVVQYGBDVVHPEHIND	149
Qy	219	YRSVKDVEAASHIEQRGGTTETRTAFGIEFARSEAFQGGKRGAKKWMIVITDGESHDSP	278
Db	150	YRSVKDVEAASHIEQRGGTTETRTAFGIEFARSEAFQGGKRGAKKWMIVITDGESHDSP	209
Qy	279	DLEKVIQOSERDNNVTRVAVAVLGYNNRRGINPEFLMEIKYIASDPDDKHFNVTDEAAL	338
Db	210	DLEKVIQOSERDNNVTRVAVAVLGYNNRRGINPEFLMEIKYIASDPDDKHFNVTDEAAL	269
Qy	339	KDVIDALGDRI FSLLEGNNKNETSPGLEMSOTGFSSSHVEDGVLGAVGAYDNGAVLKET	398
Db	270	KDVIDALGDRI FSLLEGNNKNETSPGLEMSOTGFSSSHVEDGVLGAVGAYDNGAVLKET	329
Qy	399	SAGKVIPLRESYLKEPPEELKNHGAYLGYTVTVSVSSRQGRVYVAGAPRFNHTKVLFT	458
Db	330	SAGKVIPLRESYLKEPPEELKNHGAYLGYTVTVSVSSRQGRVYVAGAPRFNHTKVLFT	389
Qy	459	MHNRSUTIHQMRGQOIGSYFGESETSVDDIGDGVTDVLLVGAPMKNFNEGRERKGVVY	518
Db	390	MHNRSUTIHQMRGQOIGSYFGESETSVDDIGDGVTDVLLVGAPMKNFNEGRERKGVVY	449
Qy	519	ELQNRVYNGTLKDSHSYQNARPGSSIASVRDLNQDSYNDVVVGAPELDNHAGAIYIFH	578
Db	450	ELQNRVYNGTLKDSHSYQNARPGSSIASVRDLNQDSYNDVVVGAPELDNHAGAIYIFH	509
Qy	579	GFRGSLTKPKORITASELATGLQYFCGSIHGQDLNEDGLIDIAVAGLGNVILMSRPV	638
Db	510	GFRGSLTKPKORITASELATGLQYFCGSIHGQDLNEDGLIDIAVAGLGNVILMSRPV	569
Qy	639	VOINASHPEPSKINIFHRQCKSGRDTCLAAFLCPTPIFLAPHFOITTVGIRYNATMD	698
Db	570	VOINASHPEPSKINIFHRQCKSGRDTCLAAFLCPTPIFLAPHFOITTVGIRYNATMD	629
Qy	699	ERRYTPRAHLDEGGDRFTNRAVILSSQGLCERINPHVLDITADTVKPTVPFVBSYLEPDP	758
Db	630	ERRYTPRAHLDEGGDRFTNRAVILSSQGLCERINPHVLDITADTVKPTVPFVBSYLEPDP	689
Qy	759	HGPMWLDGQWPTTLRVSPVPFNGCNEDEHCVPDVLVDARSDLPTAMEYCORVLRKPAQDCS	818
Db	690	HGPMWLDGQWPTTLRVSPVPFNGCNEDEHCVPDVLVDARSDLPTAMEYCORVLRKPAQDCS	749
Qy	819	AYTLSPDFTTFPIESTRQRVAVATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGS	878
Db	750	AYTLSPDFTTFPIESTRQRVAVATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGS	809
Qy	879	IECWBERRLKQVCNVSYSPFFRAKAKVAFRLD	911
Db	810	IECWBERRLKQVCNVSYSPFFRAKAKVAFRLD	842

## RESULT 9

```

US-09-764-870-313
; Sequence 313, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCES: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 313
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-313
Query Match 41.2%; Score 489; DB 9; Length 707;
Best Local Similarity 99.7%; Pred. No. 0;

```

Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MDLPRGLVAVWALSMPGFTDTFNMTRKPRVPGSRTPAFPGYTVQOHDISGNKWLAVGA 60  
 Db 17 MDLPRGLVAVWALSMPGFTDTFNMTRKPRVPGSRTPAFPGYTVQOHDISGNKWLAVGA 76

Qy 61 PLETNQYQKTDGVYKCPVHGNCTKLNLRVTLNVSERKDNMRGLSLATNPKDNSFLA 120  
 Db 77 PLETNQYQKTDGVYKCPVHGNCTKLNLRVTLNVSERKDNMRGLSLATNPKDNSFLA 136

Qy 121 CSPLWSHECGSSYYTTGMCNRSNFRFSKTVAPALQRCQTYMDIVVLDSGNSIYPWVE 180  
 Db 137 CSPLWSHECGSSYYTTGMCNRSNFRFSKTVAPALQRCQTYMDIVVLDSGNSIYPWVE 196

Qy 181 VOHFLINILKKFYIGPQIQGVVQYGEDVWHEFLNDYRSVKDVVEAASHIEQGGTET 240  
 Db 197 VOHFLINILKKFYIGPQIQGVVQYGEDVWHEFLNDYRSVKDVVEAASHIEQGGTET 256

Qy 241 RTAFGIEFARSEAPKQGRGAKKVMIVITDGHSDSPDLKVIQOESRDNVTRYAVAVL 300  
 Db 257 RTAFGIEFARSEAPKQGRGAKKVMIVITDGHSDSPDLKVIQOESRDNVTRYAVAVL 316

Qy 301 GYNNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSLGNTKNET 360  
 Db 317 GYNNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSLGNTKNET 376

Qy 361 SPGLEMSOTGFSSHVDEGVLLGAVGAYDWNGAVLKTSAKVIPLRESYLKPEPEELKN 420  
 Db 377 SPGLEMSOTGFSSHVDEGVLLGAVGAYDWNGAVLKTSAKVIPLRESYLKPEPEELKN 436

Qy 421 HGAYLGYTVTSVSSRQGRVYVAGAPRFNHTGKVILFTMHNNRSLTTHQAMRGQIQSYF 480  
 Db 437 HGAYLGYTVTSVSSRQGRVYVAGAPRFNHTGKVILFTMHNNRSLTTHQAMRGQIQSYF 496

Qy 481 GSEITSVDIDGVTDLVLLGAPMYFNEGRERKGVYVYELQNRVYNGTLKDSHYQNA 540  
 Db 497 GSEITSVDIDGVTDLVLLGAPMYFNEGRERKGVYVYELQNRVYNGTLKDSHYQNA 556

Qy 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFSGSILKTPKQITASELATG 600  
 Db 557 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFSGSILKTPKQITASELATG 616

Qy 601 LOYFGCSIHQGLDNEGLDLAVGALGNVILWSPVQVNASLHFPSPKINIFHRDCK 660  
 Db 617 LOYFGCSIHQGLDNEGLDLAVGALGNVILWSPVQVNASLHFPSPKINIFHRDCK 676

Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTNGI 691  
 Db 677 RSGRDATCLAAFLCFTPIFLAPHFQTTTNGI 707

RESULT 10  
 US-10-125-540-313  
 ; Sequence 313, Application US/10125540  
 ; Publication No. US20030059875A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: P7214C1  
 ; CURRENT APPLICATION NUMBER: US/10/125,540  
 ; CURRENT FILING DATE: 2002-04-19  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 646  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 313  
 ; LENGTH: 707  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-125-540-313

Query Match 41.2%; Score 489; DB 14; Length 707;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 11

US-10-108-260A-3386  
 ; Sequence 3386, Application US/10108260A  
 ; Publication No. US20040005560A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HELIX RESEARCH INSTITUTE  
 ; TITLE OF INVENTION: No. US20040005560A1el full length cDNA  
 ; FILE REFERENCE: H1-A0106  
 ; CURRENT APPLICATION NUMBER: US/10/108,260A  
 ; CURRENT FILING DATE: 2002-03-27  
 ; NUMBER OF SEQ ID NOS: 5458  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 3386  
 ; LENGTH: 437  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-108-260A-3386

Query Match 20.6%; Score 245; DB 15; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-225;  
 Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 175 IYPWVEVQHLINILKKFYIGPQIQGVVQYGEDVWHEFLNDYRSVKDVVEAASHIEQ 234

```

Db 175 IYFVVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDVVEAASHIEQ 234
Qy 235 RGTGTTTAFGIEFARSEAFQKGRKGAKKVMIVITGESHDSPPDLEKVIQOQSERDNVTR 294
Db 235 RGTGTTTAFGIEFARSEAFQKGRKGAKKVMIVITGESHDSPPDLEKVIQOQSERDNVTR 294
Qy 295 YAVAVLGYNNRGINPFTFLNEIKYIASDPDDKHFFNNVTDEAALKDVIDALGDRIFSLEG 354
Db 295 YAVAVLGYNNRGINPFTFLNEIKYIASDPDDKHFFNNVTDEAALKDVIDALGDRIFSLEG 354
Qy 355 TNRNETSFGLMSGTGFSHVVEDGVLGAVGAYDMNGAVLKTSAGKVIPLRESYKEF 414
Db 355 TNRNETSFGLMSGTGFSHVVEDGVLGAVGAYDMNGAVLKTSAGKVIPLRESYKEF 414
Qy 415 PEELK 419
Db 415 PEELK 419

```

```

RESULT 12
US-09-805-354-8
; Sequence 8, Application US/09805354
; Publication No. US20030078375A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-536001
; CURRENT APPLICATION NUMBER: US/09/805,354
; CURRENT FILING DATE: 2002-06-04
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-805-354-8

```

```

Query Match 16.2%; Score 193; DB 10; Length 193;
Best Local Similarity 100.0%; Pred. No. 1.9e-175;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVEHFLND 218
Db 1 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVEHFLND 60

Qy 219 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQKGRKGAKKVMIVITGESHDS 278
Db 61 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQKGRKGAKKVMIVITGESHDS 120

Qy 279 DLEKVIQOQSERDNVTRYAVAVLGYNNRGINPFTFLNEIKYIASDPDDKHFFNNVTDEAAL 338
Db 121 DLEKVIQOQSERDNVTRYAVAVLGYNNRGINPFTFLNEIKYIASDPDDKHFFNNVTDEAAL 180

Qy 339 KDVIDALGDRIFS 351
Db 181 KDVIDALGDRIFS 193

```

```

RESULT 13
US-09-758-493-8
; Sequence 8, Application US/09758493
; Publication No. US20040086935A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping

```

```

; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 00786-804001
; CURRENT APPLICATION NUMBER: US/09/758,493
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-493-8

```

```

Query Match 16.2%; Score 193; DB 12; Length 193;
Best Local Similarity 100.0%; Pred. No. 1.9e-175;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVEHFLND 218
Db 1 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVEHFLND 60

Qy 219 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQKGRKGAKKVMIVITGESHDS 278
Db 61 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQKGRKGAKKVMIVITGESHDS 120

Qy 279 DLEKVIQOQSERDNVTRYAVAVLGYNNRGINPFTFLNEIKYIASDPDDKHFFNNVTDEAAL 338
Db 121 DLEKVIQOQSERDNVTRYAVAVLGYNNRGINPFTFLNEIKYIASDPDDKHFFNNVTDEAAL 180

Qy 339 KDVIDALGDRIFS 351
Db 181 KDVIDALGDRIFS 193

```

```

RESULT 14
US-10-144-259-8
; Sequence 8, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-8

```

```

Query Match 16.2%; Score 193; DB 14; Length 193;
Best Local Similarity 100.0%; Pred. No. 1.9e-175;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVEHFLND 218
Db 1 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGPGQIQGVVQYGEDVVEHFLND 60

Qy 219 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQKGRKGAKKVMIVITGESHDS 278
Db 61 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQKGRKGAKKVMIVITGESHDS 120

Qy 279 DLEKVIQOQSERDNVTRYAVAVLGYNNRGINPFTFLNEIKYIASDPDDKHFFNNVTDEAAL 338

```

Db 121 DLEKVIQQSERDNTYAYAVLGYNRRGINPETFLNEIKYIASDPDDKHFENVIDEAL 180  
Qy 339 KOIVDALGDRIFS 351  
Db 181 KOIVDALGDRIFS 193

RESULT 15  
US-09-764-870-472  
; Sequence 472, Application US/09764870  
; Patent No. US20020042386A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT214  
; CURRENT APPLICATION NUMBER: US/09/764,870  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 646  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 472  
; LENGTH: 103  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (96)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-870-472

Query Match 7.4%; Score 88; DB 9; Length 103;  
Best Local Similarity 100.0%; Pred. No. 2.1e-75;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVWVHEFLND 218  
Db 8 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVWVHEFLND 67  
Qy 219 YRSVKDVEAAASHIEQGGTETRTAFGI 246  
Db 68 YRSVKDVEAAASHIEQGGTETRTAFGI 95

RESULT 16  
US-09-764-875-1193  
; Sequence 1193, Application US/09764875  
; Publication No. US20040018969A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P3202  
; CURRENT APPLICATION NUMBER: US/09/764,875  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1249  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1193  
; LENGTH: 103  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (96)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-875-1193

Query Match 7.4%; Score 88; DB 11; Length 103;  
Best Local Similarity 100.0%; Pred. No. 2.1e-75;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVWVHEFLND 218

Db 8 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVWVHEFLND 67  
Qy 219 YRSVKDVEAAASHIEQGGTETRTAFGI 246  
Db 68 YRSVKDVEAAASHIEQGGTETRTAFGI 95

RESULT 17  
US-10-125-540-472  
; Sequence 472, Application US/10125540  
; Publication No. US20030059875A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT214C1  
; CURRENT APPLICATION NUMBER: US/10/125,540  
; CURRENT FILING DATE: 2002-04-19  
; Prior Application removed - See File wrapper or Palm  
; NUMBER OF SEQ ID NOS: 646  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 472  
; LENGTH: 103  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (96)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-125-540-472

Query Match 7.4%; Score 88; DB 14; Length 103;  
Best Local Similarity 100.0%; Pred. No. 2.1e-75;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 159 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVWVHEFLND 218  
Db 8 CQTYMDIVIVLDGNSIYPWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVWVHEFLND 67  
Qy 219 YRSVKDVEAAASHIEQGGTETRTAFGI 246  
Db 68 YRSVKDVEAAASHIEQGGTETRTAFGI 95

RESULT 18  
US-10-144-259-25  
; Sequence 25, Application US/10144259  
; Publication No. US20030109691A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-548001  
; CURRENT APPLICATION NUMBER: US/10/144,259  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: US 09/758,493  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-144-259-25

Query Match 3.5%; Score 42; DB 14; Length 42;  
Best Local Similarity 100.0%; Pred. No. 7.9e-32;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 335 EAALKDIVDALGDRIFSELTGNTKNETSPGLEMSQTGFSSHYV 376



Db 1 EAALKDIVALGDRIFSLGKTKNETSFGLEMSQTGFSSHV 42

## RESULT 19

US-09-866-050A-500  
; Sequence 500, Application US/09866050A  
; Publication No. US20030040471A1

## GENERAL INFORMATION:

; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Murison, James G.  
; APPLICANT: Murison, James G.  
; APPLICANT: Murison, James G.  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; TITLE OF INVENTION: and Methods for Their Use  
; FILE REFERENCE: 11000.1011c4U  
; CURRENT APPLICATION NUMBER: US/09/866.050A  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 500  
; LENGTH: 545  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-866-050A-500

Query Match 3.5%; Score 42; DB 10; Length 545;  
Best Local Similarity 100.0%; Pred. No. 7.8e-31;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 GWPTTLRVSPFWNGCNEDEHCVPDLVLDARSDLPTAMEYCO 807

Db 123 GWPTTLRVSPFWNGCNEDEHCVPDLVLDARSDLPTAMEYCO 164

## RESULT 20

US-09-866-050A-624  
; Sequence 624, Application US/09866050A  
; Publication No. US20030040471A1

## GENERAL INFORMATION:

; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Murison, James G.  
; APPLICANT: Murison, James G.  
; APPLICANT: Murison, James G.  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; TITLE OF INVENTION: and Methods for Their Use  
; FILE REFERENCE: 11000.1011c4U  
; CURRENT APPLICATION NUMBER: US/09/866.050A  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 624  
; LENGTH: 688  
; TYPE: PRT  
; ORGANISM: Rat  
US-09-866-050A-624

Query Match 3.5%; Score 42; DB 10; Length 688;  
Best Local Similarity 100.0%; Pred. No. 9.6e-31;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 GWPTTLRVSPFWNGCNEDEHCVPDLVLDARSDLPTAMEYCO 807

Db 266 GWPTTLRVSPFWNGCNEDEHCVPDLVLDARSDLPTAMEYCO 307

## RESULT 21

US-09-866-050A-501  
; Sequence 501, Application US/09866050A  
; Publication No. US20030040471A1

## GENERAL INFORMATION:

; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Murison, Rene  
; APPLICANT: Murison, Rene  
; APPLICANT: Murison, James G.  
; APPLICANT: Murison, James G.  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; TITLE OF INVENTION: and Methods for Their Use  
; FILE REFERENCE: 11000.1011c4U  
; CURRENT APPLICATION NUMBER: US/09/866.050A  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 501  
; LENGTH: 696  
; TYPE: PRT  
; ORGANISM: Rat  
US-09-866-050A-501

Query Match 3.5%; Score 42; DB 10; Length 696;  
Best Local Similarity 100.0%; Pred. No. 9.7e-31;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 GWPTTLRVSPFWNGCNEDEHCVPDLVLDARSDLPTAMEYCO 807

Db 266 GWPTTLRVSPFWNGCNEDEHCVPDLVLDARSDLPTAMEYCO 307

## RESULT 22

US-09-984-130-102  
; Sequence 102, Application US/09984130  
; Publication No. US2003005231A1

## GENERAL INFORMATION:

; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 12 Human Secreted Proteins  
; FILE REFERENCE: PF489P2  
; CURRENT APPLICATION NUMBER: US/09/984.130  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/243,792  
; PRIOR FILING DATE: 2000-10-30  
; PRIOR APPLICATION NUMBER: 09/836,353  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/198,407  
; PRIOR FILING DATE: 2000-04-19  
; PRIOR APPLICATION NUMBER: PCT/US99/25031  
; PRIOR FILING DATE: 1999-10-27  
; PRIOR APPLICATION NUMBER: 60/105,971  
; PRIOR FILING DATE: 1998-10-28  
; NUMBER OF SEQ ID NOS: 149  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 102  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-984-130-102

Query Match 2.8%; Score 33; DB 10; Length 33;  
Best Local Similarity 100.0%; Pred. No. 2.5e-23;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 64 TNGYQKTGVDYKCPVTHGNCCTKLNLRVTLSNV 96

Db 1 TNGYQKTGVDYKCPVTHGNCCTKLNLRVTLSNV 33

## RESULT 23

US-09-836-353A-102  
; Sequence 102, Application US/09836353A  
; Publication No. US20030129685A1

## GENERAL INFORMATION:

; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 12 Human Secreted Proteins

```
; FILE REFERENCE: PF489P1
; CURRENT APPLICATION NUMBER: US/09/836,353A
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 102
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-353A-102

Query Match      2.8%; Score 33; DB 10; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.5e-23;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 TNGYQKTGVDYKCPVHGNCTKLNLRVTLNV 96
Db 1 TNGYQKTGVDYKCPVHGNCTKLNLRVTLNV 33

RESULT 24
US-09-866-050A-391
; Sequence 391, Application US/09866050A
; Publication No. US2003004071A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 391
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-391

Query Match      1.8%; Score 21; DB 10; Length 158;
Best Local Similarity 100.0%; Pred. No. 2.9e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 RPVVQINASHLHPEPSKINIFH 656
Db 57 RPVVQINASHLHPEPSKINIFH 77

RESULT 25
US-09-866-050A-390
; Sequence 390, Application US/09866050A
; Publication No. US2003004071A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
```

```
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 390
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-390

Query Match      1.4%; Score 17; DB 10; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 NGTLKDSHSYQNAFPGS 544
Db 28 NGTLKDSHSYQNAFPGS 44

RESULT 26
US-09-764-870-284
; Sequence 284, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 284
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-284

Query Match      1.3%; Score 15; DB 9; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYFW 178
Db 47 DIVIVLDGNSIYFW 61

RESULT 27
US-09-764-870-444
; Sequence 444, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 444
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-444

Query Match      1.3%; Score 15; DB 9; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYFW 178
Db 47 DIVIVLDGNSIYFW 61
```

## RESULT 28

US-10-125-540-284  
; Sequence 284, Application US/10125540  
; Publication No. US20030059875A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT214C1  
; CURRENT APPLICATION NUMBER: US/10/125,540  
; CURRENT FILING DATE: 2002-04-19  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 646  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 284  
; LENGTH: 148  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-125-540-284

Query Match 1.3%; Score 15; DB 14; Length 148;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGSGNSIYPW 178  
Db 47 DIVVLGSGNSIYPW 61  
|||||

## RESULT 29

US-10-125-540-444  
; Sequence 444, Application US/10125540  
; Publication No. US20030059875A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT214C1  
; CURRENT APPLICATION NUMBER: US/10/125,540  
; CURRENT FILING DATE: 2002-04-19  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 646  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 444  
; LENGTH: 148  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-125-540-444

Query Match 1.3%; Score 15; DB 14; Length 148;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGSGNSIYPW 178  
Db 47 DIVVLGSGNSIYPW 61  
|||||

## RESULT 30

US-10-346-863-57  
; Sequence 57, Application US/10346863  
; Publication No. US20040038325A1  
; GENERAL INFORMATION:  
; APPLICANT: PHELPS, CHRISTOPHER BENJAMIN  
; APPLICANT: FAGAN, RICHARD JOSEPH  
; APPLICANT: GUTTERIDGE, ALEX  
; TITLE OF INVENTION: ADHESION MOLECULES  
; FILE REFERENCE: 674575-2001  
; CURRENT APPLICATION NUMBER: US/10/346,863  
; CURRENT FILING DATE: 2003-01-17  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 646  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 57  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-346-863-57

Query Match 1.3%; Score 15; DB 12; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGSGNSIYPW 178  
Db 5 DIVVLGSGNSIYPW 19  
|||||

## RESULT 31

US-10-474-832-59  
; Sequence 59, Application US/10474832  
; Publication No. US20040081651A1  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; TITLE OF INVENTION: ANTIBODIES TO VLA-1  
; FILE REFERENCE: A101 PCT  
; CURRENT APPLICATION NUMBER: US/10/474,832  
; CURRENT FILING DATE: 2003-10-14  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 60/283,794  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 59  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Rat/human chimeric  
; OTHER INFORMATION: I domain construct  
US-10-474-832-59

Query Match 1.3%; Score 15; DB 16; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGSGNSIYPW 178  
Db 4 DIVVLGSGNSIYPW 18  
|||||

## RESULT 32

US-10-474-832-60  
; Sequence 60, Application US/10474832  
; Publication No. US20040081651A1  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; TITLE OF INVENTION: ANTIBODIES TO VLA-1  
; FILE REFERENCE: A101 PCT  
; CURRENT APPLICATION NUMBER: US/10/474,832  
; CURRENT FILING DATE: 2003-10-14  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 60/283,794  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 60  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Rat/human chimeric  
; OTHER INFORMATION: I domain construct  
US-10-474-832-60

Query Match 1.3%; Score 15; DB 12; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGSGNSIYPW 178  
Db 5 DIVVLGSGNSIYPW 19  
|||||

; ORGANISM: Rattus sp.  
US-10-474-832-60

Query Match 1.3%; Score 15; DB 16; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGDSNSIYPW 178  
|||||  
DB 4 DIVVLGDSNSIYPW 18

RESULT 33

US-10-474-832-61  
; Sequence 61, Application US/10474832  
; Publication No. US20040081651A1  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; TITLE OF INVENTION: ANTIBODIES TO VLA-1  
; FILE REFERENCE: AL01 PCT  
; CURRENT APPLICATION NUMBER: US/10/474,832  
; CURRENT FILING DATE: 2003-10-14  
; PRIOR APPLICATION NUMBER: 60/283,794  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: 60/303,689  
; PRIOR FILING DATE: 2001-07-06  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 61  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-474-832-61

Query Match 1.3%; Score 15; DB 16; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGDSNSIYPW 178  
|||||  
DB 4 DIVVLGDSNSIYPW 18

RESULT 34

US-09-805-354-5  
; Sequence 5, Application US/09805354  
; Publication No. US20030078375A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Rui  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-536001  
; CURRENT APPLICATION NUMBER: US/09/805,354  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: US 09/758,493  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-805-354-5

Query Match 1.3%; Score 15; DB 10; Length 195;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGDSNSIYPW 178  
|||||

Db 6 DIVVLGDSNSIYPW 20

RESULT 35

US-09-758-493-5  
; Sequence 5, Application US/09758493  
; Publication No. US20040086935A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND  
; FILE REFERENCE: 00786-804001  
; CURRENT APPLICATION NUMBER: US/09/758,493  
; CURRENT FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-758-493-5

Query Match 1.3%; Score 15; DB 12; Length 195;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGDSNSIYPW 178  
|||||  
DB 6 DIVVLGDSNSIYPW 20

RESULT 36

US-10-144-259-5  
; Sequence 5, Application US/10144259  
; Publication No. US20030109691A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-548001  
; CURRENT APPLICATION NUMBER: US/10/144,259  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: US 09/758,493  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-144-259-5

Query Match 1.3%; Score 15; DB 14; Length 195;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGDSNSIYPW 178  
|||||  
DB 6 DIVVLGDSNSIYPW 20

RESULT 37

US-09-996-738-5  
; Sequence 5, Application US/09996738  
; Patent No. US20020146417A1  
; GENERAL INFORMATION:

```
; APPLICANT: Biogen, Inc.
; APPLICANT: De Fougerolles, Antonin
; APPLICANT: Gotwals, Philip
; APPLICANT: Lobb, Roy
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Inflammatory
; TITLE OF INVENTION: Disorders
; FILE REFERENCE: A076PCT
; CURRENT APPLICATION NUMBER: US/09/996,738
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/185336
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 60/137038
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 212
; TYPE: PRT
; ORGANISM: rat
US-09-996-738-5
```

```
Query Match 1.3%; Score 15; DB 9; Length 212;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 164 DIVVLGSGNSIYPW 178
Db 22 DIVVLGSGNSIYPW 36
```

## RESULT 38

```
US-09-996-738-5
; Sequence 6, Application US/09996738
; Patent No. US20020146417A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: De Fougerolles, Antonin
; APPLICANT: Gotwals, Philip
; APPLICANT: Lobb, Roy
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Inflammatory
; TITLE OF INVENTION: Disorders
; FILE REFERENCE: A076PCT
; CURRENT APPLICATION NUMBER: US/09/996,738
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/185336
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 60/137038
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-996-738-6
```

```
Query Match 1.3%; Score 15; DB 9; Length 212;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 164 DIVVLGSGNSIYPW 178
Db 22 DIVVLGSGNSIYPW 36
```

## RESULT 39

```
US-10-625-260-5
; Sequence 5, Application US/10625260
; Publication No. US20040037827A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
```

```
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/625,260
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Rat
US-10-625-260-5
```

```
Query Match 1.3%; Score 15; DB 12; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 164 DIVVLGSGNSIYPW 178
Db 22 DIVVLGSGNSIYPW 36
```

## RESULT 40

```
US-10-625-260-6
; Sequence 6, Application US/10625260
; Publication No. US20040037827A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/625,260
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-625-260-6
```

```
Query Match 1.3%; Score 15; DB 12; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 164 DIVVLGSGNSIYPW 178
Db 22 DIVVLGSGNSIYPW 36
```

## RESULT 41

```
US-10-625-260-9
; Sequence 9, Application US/10625260
; Publication No. US20040037827A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/625,260
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: 60/137,214
```

; PRIOR FILING DATE: 1999-06-01  
 ; PRIOR APPLICATION NUMBER: 60/130,847  
 ; PRIOR FILING DATE: 1999-04-22  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PastSeq for Windows Version 4.0  
 ; SEQ ID NO 9  
 ; LENGTH: 214  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-10-625-260-9

Query Match 1.3%; Score 15; DB 12; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178  
 Db 22 DIVVLGSGNSIYPW 36

## RESULT 42

US-10-061-658-5  
 ; Sequence 5, Application US/10061658  
 ; Publication No. US20020182213A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Biogen, Inc.  
 ; APPLICANT: Gotwals, Philip  
 ; APPLICANT: Koteliarsky, Victor  
 ; TITLE OF INVENTION: Method for the Treatment of Fibrosis  
 ; FILE REFERENCE: A07305  
 ; CURRENT APPLICATION NUMBER: US/10/061,658  
 ; CURRENT FILING DATE: 2002-02-01  
 ; PRIOR APPLICATION NUMBERS: 60/137,214  
 ; PRIOR FILING DATE: 1999-06-01  
 ; PRIOR APPLICATION NUMBER: 60/130,847  
 ; PRIOR FILING DATE: 1999-04-22  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PastSeq for Windows Version 4.0  
 ; SEQ ID NO 5  
 ; LENGTH: 214  
 ; TYPE: PRT  
 ; ORGANISM: Rat  
 US-10-061-658-5

Query Match 1.3%; Score 15; DB 13; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178  
 Db 22 DIVVLGSGNSIYPW 36

## RESULT 43

US-10-061-658-6  
 ; Sequence 6, Application US/10061658  
 ; Publication No. US20020182213A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Biogen, Inc.  
 ; APPLICANT: Gotwals, Philip  
 ; APPLICANT: Koteliarsky, Victor  
 ; TITLE OF INVENTION: Method for the Treatment of Fibrosis  
 ; FILE REFERENCE: A07305  
 ; CURRENT APPLICATION NUMBER: US/10/061,658  
 ; CURRENT FILING DATE: 2002-02-01  
 ; PRIOR APPLICATION NUMBERS: 60/137,214  
 ; PRIOR FILING DATE: 1999-06-01  
 ; PRIOR APPLICATION NUMBER: 60/130,847  
 ; PRIOR FILING DATE: 1999-04-22  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PastSeq for Windows Version 4.0  
 ; SEQ ID NO 6  
 ; LENGTH: 214

; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-10-061-658-6

Query Match 1.3%; Score 15; DB 13; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178  
 Db 22 DIVVLGSGNSIYPW 36

## RESULT 44

US-10-061-658-9  
 ; Sequence 9, Application US/10061658  
 ; Publication No. US20020182213A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Biogen, Inc.  
 ; APPLICANT: Gotwals, Philip  
 ; APPLICANT: Koteliarsky, Victor  
 ; TITLE OF INVENTION: Method for the Treatment of Fibrosis  
 ; FILE REFERENCE: A07305  
 ; CURRENT APPLICATION NUMBER: US/10/061,658  
 ; CURRENT FILING DATE: 2002-02-01  
 ; PRIOR APPLICATION NUMBERS: 60/137,214  
 ; PRIOR FILING DATE: 1999-06-01  
 ; PRIOR APPLICATION NUMBER: 60/130,847  
 ; PRIOR FILING DATE: 1999-04-22  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PastSeq for Windows Version 4.0  
 ; SEQ ID NO 9  
 ; LENGTH: 214  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-10-061-658-9

Query Match 1.3%; Score 15; DB 13; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178  
 Db 22 DIVVLGSGNSIYPW 36

## RESULT 45

US-10-474-832-63  
 ; Sequence 63, Application US/10474832  
 ; Publication No. US20040081651A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BIOGEN, INC.  
 ; TITLE OF INVENTION: ANTIBODIES TO VLA-1  
 ; FILE REFERENCE: A101 PCT  
 ; CURRENT APPLICATION NUMBER: US/10/474,832  
 ; CURRENT FILING DATE: 2003-10-14  
 ; PRIOR APPLICATION NUMBERS: 60/283,794  
 ; PRIOR FILING DATE: 2001-04-13  
 ; PRIOR APPLICATION NUMBER: 60/303,689  
 ; PRIOR FILING DATE: 2001-07-06  
 ; NUMBER OF SEQ ID NOS: 70  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 63  
 ; LENGTH: 214  
 ; TYPE: PRT  
 ; ORGANISM: Rattus sp.  
 US-10-474-832-63

Query Match 1.3%; Score 15; DB 16; Length 214;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178

```
Db      22 DIVVLGSGNSIYPW 36
|||||
RESULT 46
US-10-474-832-64
; Sequence 64, Application US/10474832
; Publication No. US20040081651A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: ANTIBODIES TO VLA-1
; FILE REFERENCE: A101 PCT
; CURRENT APPLICATION NUMBER: US/10/474,832
; CURRENT FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/283,794
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/303,689
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-474-832-64

Query Match      1.3%; Score 15; DB 16; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVVLGSGNSIYPW 178
|||||
Db      22 DIVVLGSGNSIYPW 36
|||||

RESULT 47
US-09-984-130-103
; Sequence 103, Application US/09984130
; Publication No. US20030055231A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: P489P2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 103
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-103

Query Match      1.3%; Score 15; DB 10; Length 1151;
Best Local Similarity 100.0%; Pred. No. 9.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVVLGSGNSIYPW 178
|||||
Db      144 DIVVLGSGNSIYPW 158
|||||

RESULT 48
US-09-984-130-103
; Sequence 103, Application US/09836353A
; Publication No. US20030129685A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: P489P1
; CURRENT APPLICATION NUMBER: US/09/836,353A
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 103
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-353A-103

Query Match      1.3%; Score 15; DB 10; Length 1151;
Best Local Similarity 100.0%; Pred. No. 9.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVVLGSGNSIYPW 178
|||||
Db      144 DIVVLGSGNSIYPW 158
|||||

RESULT 49
US-09-918-715-250
; Sequence 250, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Brad St. Croix
; APPLICANT: Bert Vogelstein
; APPLICANT: Kenneth Kinzler
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00134
; CURRENT APPLICATION NUMBER: US/09/918,715
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/222,599
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/224,360
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 250
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-918-715-250

Query Match      1.3%; Score 15; DB 12; Length 1179;
Best Local Similarity 100.0%; Pred. No. 9.5e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVVLGSGNSIYPW 178
|||||
Db      172 DIVVLGSGNSIYPW 186
|||||

RESULT 50
US-09-918-715-307
; Sequence 307, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Brad St. Croix
```

; APPLICANT: Bert Vogelstein  
; APPLICANT: Kenneth Kinzler  
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS  
; FILE REFERENCE: 1107.00134  
; CURRENT APPLICATION NUMBER: US/09/918,715  
; CURRENT FILING DATE: 2001-08-01  
; PRIOR APPLICATION NUMBER: 60/222,599  
; PRIOR FILING DATE: 2000-08-02  
; PRIOR APPLICATION NUMBER: 60/224,360  
; PRIOR FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: 60/282,850  
; PRIOR FILING DATE: 2000-04-11  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 307  
; LENGTH: 1180  
; TYPE: PRT  
; ORGANISM: Rat  
US-09-918-715-307

Query Match 1.1%; Score 15; DB 12; Length 1180;  
Best Local Similarity 100.0%; Pred. No. 9.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178  
DB 172 DIVVLGSGNSIYPW 186  
|||||

RESULT 51  
US-09-805-354-7  
; Sequence 7, Application US/09805354  
; Publication No. US20030078375A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-536001  
; CURRENT APPLICATION NUMBER: US/09/805,354  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: US 09/758,493  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-805-354-7

Query Match 1.1%; Score 13; DB 10; Length 195;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 VIVLDGNSIYPW 178  
DB 8 VIVLDGNSIYPW 20  
|||||

RESULT 52  
US-09-758-493-7  
; Sequence 7, Application US/09758493  
; Publication No. US20040086935A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND  
; USES THEREOF  
; FILE REFERENCE: 00786-804001

; CURRENT APPLICATION NUMBER: US/09/758,493  
; CURRENT FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-758-493-7

Query Match 1.1%; Score 13; DB 12; Length 195;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 VIVLDGNSIYPW 178  
DB 8 VIVLDGNSIYPW 20  
|||||

RESULT 53  
US-10-144-259-7  
; Sequence 7, Application US/10144259  
; Publication No. US20030109691A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-548001  
; CURRENT APPLICATION NUMBER: US/10/144,259  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: US 09/758,493  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-144-259-7

Query Match 1.1%; Score 13; DB 14; Length 195;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 VIVLDGNSIYPW 178  
DB 8 VIVLDGNSIYPW 20  
|||||

RESULT 54  
US-10-144-259-26  
; Sequence 26, Application US/10144259  
; Publication No. US20030109691A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-548001  
; CURRENT APPLICATION NUMBER: US/10/144,259  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: US 09/758,493  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 26



```
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-26

Query Match          0.9%; Score 11; DB 14; Length 43;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  340 DIVDALGORIF 350
    |||||
Db   6 DIVDALGDRIF 16

RESULT 55
US-10-336-603A-42
; Sequence 42, Application US/10336603A
; Publication No. US20040072997A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-533A
; CURRENT APPLICATION NUMBER: US/10/336,603A
; CURRENT FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: 09/746,491
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 10/055,569
; PRIOR FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 42
; LENGTH: 1147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-336-603A-42

Query Match          0.9%; Score 11; DB 12; Length 1147;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  495 TDVLLVGAPMY 505
    |||||
Db   506 TDVLLVGAPMY 516

RESULT 56
US-10-211-462-187
; Sequence 187, Application US/10211462
; Publication No. US20040033495A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 187
; LENGTH: 1181
; TYPE: PRT

; ORGANISM: Homo sapiens
US-10-211-462-187

Query Match          0.9%; Score 11; DB 12; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  495 TDVLLVGAPMY 505
    |||||
Db   506 TDVLLVGAPMY 516

RESULT 57
US-10-160-354-2
; Sequence 2, Application US/10160354
; Publication No. US20030013107A1
; GENERAL INFORMATION:
; APPLICANT: Chu, Peter
; APPLICANT: Li, Congfen
; APPLICANT: Liao, X. Charlene
; APPLICANT: Pardo, Jorge
; APPLICANT: Rigel Pharmaceuticals, Incorporated
; TITLE OF INVENTION: Alpha 2 Intergrin: Modulators of Lymphocyte Activation
; FILE REFERENCE: 021044-001110US
; CURRENT APPLICATION NUMBER: US/10/160,354
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: US 60/296,819
; PRIOR FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1181
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-160-354-2

Query Match          0.9%; Score 11; DB 14; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  495 TDVLLVGAPMY 505
    |||||
Db   506 TDVLLVGAPMY 516

RESULT 58
US-10-295-027-1286
; Sequence 1286, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
```

; PRIOR APPLICATION NUMBER: US 60/332,464  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 60/334,393  
; PRIOR FILING DATE: 2001-11-29  
; PRIOR APPLICATION NUMBER: US 60/340,376  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: US 60/347,211  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 60/347,349  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/355,250  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: US 60/356,714  
; PRIOR FILING DATE: 2002-02-13  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1386  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1286  
; LENGTH: 1181  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-295-027-1286

Query Match 0.8%; Score 11; DB 15; Length 1181;  
Best Local Similarity 100.0%; Pred. No. 0.63; Mismatches 0; Indels 0; Gaps 0;  
Matches 11; Conservative 0;

Qy 495 TDVLLVGAPMY 505  
|||  
Db 506 TDVLLVGAPMY 516

## RESULT 59

US-10-108-260A-3415  
; Sequence 3415, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA  
; FILE REFERENCE: HI-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3415  
; LENGTH: 164  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-3415

Query Match 0.8%; Score 9; DB 15; Length 164;  
Best Local Similarity 100.0%; Pred. No. 8.8; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

Qy 1153 LLLALLVL 1161  
|||  
Db 99 LLLALLVL 107

## RESULT 60

US-10-038-854-102  
; Sequence 102, Application US/10038854  
; Publication No. US20040022781A1  
; GENERAL INFORMATION:  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Li, Li  
; APPLICANT: Wolenc, Adam R  
; APPLICANT: Vernet, Corine  
; APPLICANT: Eisen, Andrew J  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Tchernev, Velizar

; APPLICANT: Spaderna, Steven K  
; APPLICANT: Gorman, Linda  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Gusev, Vladimir Y  
; APPLICANT: Gangolli, Esma A  
; APPLICANT: Guo, Xiaojia S  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Bológ, Ferenc  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Gunther, Erik  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: MacDougall, John R  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-230  
; CURRENT APPLICATION NUMBER: US/10/038,854  
; CURRENT FILING DATE: 2003-01-22  
; PRIOR APPLICATION NUMBER: 60/258,928  
; PRIOR FILING DATE: 2000-12-29  
; PRIOR APPLICATION NUMBER: 60/259,415  
; PRIOR FILING DATE: 2001-01-02  
; PRIOR APPLICATION NUMBER: 60/259,785  
; PRIOR FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: 60/269,814  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 60/279,832  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/279,833  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/279,863  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/283,889  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: 60/284,447  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/286,683  
; PRIOR FILING DATE: 2001-04-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 411  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 102  
; LENGTH: 312  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-10-038-854-102

Query Match 0.8%; Score 9; DB 16; Length 312;  
Best Local Similarity 100.0%; Pred. No. 16; Mismatches 9; Conservative 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1162  
|||  
Db 3 LLLALLVL 11

## RESULT 61

US-10-369-493-2470  
; Sequence 2470, Application US/10369493  
; Publication No. US2003023675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2470  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Schizosaccharomyces pombe  
US-10-369-493-2470

Query Match 0.8%; Score 9; DB 15; Length 371;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 DGVLLGAVG 386  
Db 69 DGVLLGAVG 77  
|||||

## RESULT 62

US-09-350-259-20  
; Sequence 20, Application US/09350259  
; Patent No. US20020062008A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. US20020062008A1el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350,259  
; CURRENT FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: 09/193,043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: dog  
US-09-350-259-20

Query Match 0.7%; Score 8; DB 9; Length 17;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
Db 1 LVVGAPLE 8  
|||||

## RESULT 63

US-09-891-943-20  
; Sequence 20, Application US/09891943  
; Publication No. US2003007278A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. US2003007278A1el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/891,943  
; CURRENT FILING DATE: 2001-05-26  
; PRIOR APPLICATION NUMBER: 09/193,043  
; PRIOR FILING DATE: 1998-11-16  
; PRIOR APPLICATION NUMBER: 08/286,889  
; PRIOR FILING DATE: 1994-08-05

; PRIOR APPLICATION NUMBER: 08/362,652  
; PRIOR FILING DATE: 1994-12-21  
; PRIOR APPLICATION NUMBER: 08/943,363  
; PRIOR FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: dog  
US-09-891-943-20

Query Match 0.7%; Score 8; DB 10; Length 17;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63  
Db 1 LVVGAPLE 8  
|||||

## RESULT 64

US-10-144-259-27  
; Sequence 27, Application US/10144259  
; Publication No. US20030109691A1  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; APPLICANT: Li, Rui  
; APPLICANT: Xiong, Jian-Ping  
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF  
; FILE REFERENCE: 00786-548001  
; CURRENT APPLICATION NUMBER: US/10/144,259  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: US 09/758,493  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: US 60/221,950  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 43  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-144-259-27

Query Match 0.7%; Score 8; DB 14; Length 43;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 365 EMSQTGFS 372  
Db 32 EMSQTGFS 39  
|||||

## RESULT 65

US-09-978-295A-85  
; Sequence 85, Application US/09978295A  
; Patent No. US20020156006A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnovers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gertsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gurney, Austin L.  
 APPLICANT: Hillan, Kenneth J.  
 APPLICANT: Kijavin, Ivar J.  
 APPLICANT: Kuo, Sophia S.  
 APPLICANT: Napier, Mary A.  
 APPLICANT: Pan, James.  
 APPLICANT: Paoni, Nicholas F.  
 APPLICANT: Roy, Margaret Ann  
 APPLICANT: Shelton, David L.  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Williams, P. Micky  
 APPLICANT: Wood, William I.  
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 FILE REFERENCE: P2630PIC11  
 CURRENT APPLICATION NUMBER: US/09/978,295A  
 CURRENT FILING DATE: 2001-10-15  
 PRIOR FILING DATE: 2001-07-30  
 PRIOR APPLICATION NUMBER: 09/518585  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/062250  
 PRIOR FILING DATE: 1997-11-03  
 PRIOR APPLICATION NUMBER: 60/064249  
 PRIOR FILING DATE: 1997-11-13  
 PRIOR APPLICATION NUMBER: 60/065311  
 PRIOR FILING DATE: 1997-11-21  
 PRIOR APPLICATION NUMBER: 60/066364  
 PRIOR FILING DATE: 1998-03-10  
 PRIOR APPLICATION NUMBER: 60/077450  
 PRIOR FILING DATE: 1998-03-11  
 PRIOR APPLICATION NUMBER: 60/077632  
 PRIOR FILING DATE: 1998-03-11  
 PRIOR APPLICATION NUMBER: 60/077641  
 PRIOR FILING DATE: 1998-03-11  
 PRIOR APPLICATION NUMBER: 60/077649  
 PRIOR FILING DATE: 1998-03-11  
 PRIOR APPLICATION NUMBER: 60/077791  
 PRIOR FILING DATE: 1998-03-12  
 PRIOR APPLICATION NUMBER: 60/078004  
 PRIOR FILING DATE: 1998-03-13  
 PRIOR APPLICATION NUMBER: 60/078886  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/078936  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/078939  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/079294  
 PRIOR FILING DATE: 1998-03-25  
 PRIOR APPLICATION NUMBER: 60/079656  
 PRIOR FILING DATE: 1998-03-26  
 PRIOR APPLICATION NUMBER: 60/079664  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079689  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079663  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079728  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079786  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079920  
 PRIOR FILING DATE: 1998-03-30  
 PRIOR APPLICATION NUMBER: 60/079923  
 PRIOR FILING DATE: 1998-03-30  
 PRIOR APPLICATION NUMBER: 60/080105  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080107  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080165  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080194

PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080327  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080328  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080333  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080334  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/081070  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081049  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081071  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081195  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081203  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081229  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081955  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081817  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081819  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081952  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081838  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/082568  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082569  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082704  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082804  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082700  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082797  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082796  
 PRIOR FILING DATE: 1998-04-23  
 PRIOR APPLICATION NUMBER: 60/083336  
 PRIOR FILING DATE: 1998-04-27  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/083392  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083495  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083496  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083499  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083545  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083554  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083558  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083559  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083500  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083742  
 PRIOR FILING DATE: 1998-04-30  
 PRIOR APPLICATION NUMBER: 60/084366  
 PRIOR FILING DATE: 1998-05-05  
 PRIOR APPLICATION NUMBER: 60/084414  
 PRIOR FILING DATE: 1998-05-06

; PRIOR APPLICATION NUMBER: 60/084441  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: 60/084637  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084639  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084640  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084598  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084600  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084627  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084643  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/085339  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085338  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085323  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085582  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085700  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085689  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085579  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085580  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085573  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085704  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 9; Length 67;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLIALL 1159  
 Db 12 GLLIALL 19

RESULT 66  
 US-09-978-697-85  
 ; Sequence 85, Application US/09978697  
 ; Patent No. US2002016928A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Baker Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Ferraro, Napoleon  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Hillan, Kenneth J.  
 ; APPLICANT: Kijavini, Ivar J.  
 ; APPLICANT: Kuo, Sophia S.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James;  
 ; APPLICANT: Paoni, Nicholas P.

; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Shelton, David L.  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; FILE REFERENCE: P2630PIC27  
 ; CURRENT APPLICATION NUMBER: US/09/978.697  
 ; CURRENT FILING DATE: 2001-10-16  
 ; PRIOR APPLICATION NUMBER: 09/918585  
 ; PRIOR FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: 60/062250  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/064249  
 ; PRIOR FILING DATE: 1997-11-03  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066364  
 ; PRIOR FILING DATE: 1997-11-21  
 ; PRIOR APPLICATION NUMBER: 60/077450  
 ; PRIOR FILING DATE: 1998-03-10  
 ; PRIOR APPLICATION NUMBER: 60/077632  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077641  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077649  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077791  
 ; PRIOR FILING DATE: 1998-03-12  
 ; PRIOR APPLICATION NUMBER: 60/078004  
 ; PRIOR FILING DATE: 1998-03-13  
 ; PRIOR APPLICATION NUMBER: 60/078886  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078936  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078939  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/079294  
 ; PRIOR FILING DATE: 1998-03-25  
 ; PRIOR APPLICATION NUMBER: 60/079656  
 ; PRIOR FILING DATE: 1998-03-26  
 ; PRIOR APPLICATION NUMBER: 60/079664  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079689  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079663  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079728  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079786  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079920  
 ; PRIOR FILING DATE: 1998-03-30  
 ; PRIOR APPLICATION NUMBER: 60/079923  
 ; PRIOR FILING DATE: 1998-03-30  
 ; PRIOR APPLICATION NUMBER: 60/080105  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080107  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080165  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080194  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080327  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080328  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080333  
 ; PRIOR FILING DATE: 1998-04-01

;; PRIOR APPLICATION NUMBER: 60/080334  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/081070  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081049  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081071  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081195  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081955  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081819  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081952  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081838  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082568  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082569  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082704  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082804  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082700  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082797  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082796  
;; PRIOR FILING DATE: 1998-04-23  
;; PRIOR APPLICATION NUMBER: 60/083336  
;; PRIOR FILING DATE: 1998-04-27  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083392  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083495  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083496  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083499  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083554  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083558  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083559  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083500  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083742  
;; PRIOR FILING DATE: 1998-04-30  
;; PRIOR APPLICATION NUMBER: 60/084366  
;; PRIOR FILING DATE: 1998-05-05  
;; PRIOR APPLICATION NUMBER: 60/084414  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084441  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084639  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084640

;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084598  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 9; Length 67;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159  
Db 12 GLLLLALL 19

RESULT 67  
US-09-978-192A-85  
; Sequence 85, Application US/09978192A  
; Patent No. US2002017753A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Quang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Garney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kijavini, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

1 ; TITLE OF INVENTION: Acids Encoding the Same  
2 ; FILE REFERENCE: 22630PIC9  
3 ; CURRENT APPLICATION NUMBER: US/09/978,192A  
4 ; CURRENT FILING DATE: 2001-10-15  
5 ; PRIOR APPLICATION NUMBER: 09/918585  
6 ; PRIOR FILING DATE: 2001-07-30  
7 ; PRIOR APPLICATION NUMBER: 60/062250  
8 ; PRIOR FILING DATE: 1997-10-17  
9 ; PRIOR APPLICATION NUMBER: 60/064249  
10 ; PRIOR FILING DATE: 1997-11-03  
11 ; PRIOR APPLICATION NUMBER: 60/065311  
12 ; PRIOR FILING DATE: 1997-11-13  
13 ; PRIOR APPLICATION NUMBER: 60/066364  
14 ; PRIOR FILING DATE: 1997-11-21  
15 ; PRIOR APPLICATION NUMBER: 60/077450  
16 ; PRIOR FILING DATE: 1998-03-10  
17 ; PRIOR APPLICATION NUMBER: 60/077632  
18 ; PRIOR FILING DATE: 1998-03-11  
19 ; PRIOR APPLICATION NUMBER: 60/077641  
20 ; PRIOR FILING DATE: 1998-03-11  
21 ; PRIOR APPLICATION NUMBER: 60/077649  
22 ; PRIOR FILING DATE: 1998-03-11  
23 ; PRIOR APPLICATION NUMBER: 60/077791  
24 ; PRIOR FILING DATE: 1998-03-12  
25 ; PRIOR APPLICATION NUMBER: 60/078004  
26 ; PRIOR FILING DATE: 1998-03-13  
27 ; PRIOR APPLICATION NUMBER: 60/078886  
28 ; PRIOR FILING DATE: 1998-03-20  
29 ; PRIOR APPLICATION NUMBER: 60/078936  
30 ; PRIOR FILING DATE: 1998-03-20  
31 ; PRIOR APPLICATION NUMBER: 60/078910  
32 ; PRIOR FILING DATE: 1998-03-20  
33 ; PRIOR APPLICATION NUMBER: 60/078939  
34 ; PRIOR FILING DATE: 1998-03-20  
35 ; PRIOR APPLICATION NUMBER: 60/079294  
36 ; PRIOR FILING DATE: 1998-03-25  
37 ; PRIOR APPLICATION NUMBER: 60/079656  
38 ; PRIOR FILING DATE: 1998-03-26  
39 ; PRIOR APPLICATION NUMBER: 60/079664  
40 ; PRIOR FILING DATE: 1998-03-27  
41 ; PRIOR APPLICATION NUMBER: 60/079689  
42 ; PRIOR FILING DATE: 1998-03-27  
43 ; PRIOR APPLICATION NUMBER: 60/079663  
44 ; PRIOR FILING DATE: 1998-03-27  
45 ; PRIOR APPLICATION NUMBER: 60/079728  
46 ; PRIOR FILING DATE: 1998-03-27  
47 ; PRIOR APPLICATION NUMBER: 60/079786  
48 ; PRIOR FILING DATE: 1998-03-27  
49 ; PRIOR APPLICATION NUMBER: 60/079920  
50 ; PRIOR FILING DATE: 1998-03-30  
51 ; PRIOR APPLICATION NUMBER: 60/079923  
52 ; PRIOR FILING DATE: 1998-03-30  
53 ; PRIOR APPLICATION NUMBER: 60/080105  
54 ; PRIOR FILING DATE: 1998-03-31  
55 ; PRIOR APPLICATION NUMBER: 60/080107  
56 ; PRIOR FILING DATE: 1998-03-31  
57 ; PRIOR APPLICATION NUMBER: 60/080165  
58 ; PRIOR FILING DATE: 1998-03-31  
59 ; PRIOR APPLICATION NUMBER: 60/080194  
60 ; PRIOR FILING DATE: 1998-03-31  
61 ; PRIOR APPLICATION NUMBER: 60/080327  
62 ; PRIOR FILING DATE: 1998-04-01  
63 ; PRIOR APPLICATION NUMBER: 60/080328  
64 ; PRIOR FILING DATE: 1998-04-01  
65 ; PRIOR APPLICATION NUMBER: 60/080333  
66 ; PRIOR FILING DATE: 1998-04-01  
67 ; PRIOR APPLICATION NUMBER: 60/080334  
68 ; PRIOR FILING DATE: 1998-04-01  
69 ; PRIOR APPLICATION NUMBER: 60/081070  
70 ; PRIOR FILING DATE: 1998-04-08  
71 ; PRIOR APPLICATION NUMBER: 60/081049  
72 ; PRIOR FILING DATE: 1998-04-08  
73 ; PRIOR APPLICATION NUMBER: 60/081071  
74 ; PRIOR FILING DATE: 1998-04-08  
75 ; PRIOR APPLICATION NUMBER: 60/081195  
76 ; PRIOR FILING DATE: 1998-04-08  
77 ; PRIOR APPLICATION NUMBER: 60/081203  
78 ; PRIOR FILING DATE: 1998-04-09  
79 ; PRIOR APPLICATION NUMBER: 60/081229  
80 ; PRIOR FILING DATE: 1998-04-09  
81 ; PRIOR APPLICATION NUMBER: 60/081955  
82 ; PRIOR FILING DATE: 1998-04-15  
83 ; PRIOR APPLICATION NUMBER: 60/081817  
84 ; PRIOR FILING DATE: 1998-04-15  
85 ; PRIOR APPLICATION NUMBER: 60/081819  
86 ; PRIOR FILING DATE: 1998-04-15  
87 ; PRIOR APPLICATION NUMBER: 60/081952  
88 ; PRIOR FILING DATE: 1998-04-15  
89 ; PRIOR APPLICATION NUMBER: 60/081838  
90 ; PRIOR FILING DATE: 1998-04-15  
91 ; PRIOR APPLICATION NUMBER: 60/082568  
92 ; PRIOR FILING DATE: 1998-04-21  
93 ; PRIOR APPLICATION NUMBER: 60/082569  
94 ; PRIOR FILING DATE: 1998-04-21  
95 ; PRIOR APPLICATION NUMBER: 60/082704  
96 ; PRIOR FILING DATE: 1998-04-22  
97 ; PRIOR APPLICATION NUMBER: 60/082804  
98 ; PRIOR FILING DATE: 1998-04-22  
99 ; PRIOR APPLICATION NUMBER: 60/082700  
100 ; PRIOR FILING DATE: 1998-04-22  
101 ; PRIOR APPLICATION NUMBER: 60/082797  
102 ; PRIOR FILING DATE: 1998-04-22  
103 ; PRIOR APPLICATION NUMBER: 60/082796  
104 ; PRIOR FILING DATE: 1998-04-23  
105 ; PRIOR APPLICATION NUMBER: 60/083336  
106 ; PRIOR FILING DATE: 1998-04-27  
107 ; PRIOR APPLICATION NUMBER: 60/083322  
108 ; PRIOR FILING DATE: 1998-04-28  
109 ; PRIOR APPLICATION NUMBER: 60/083392  
110 ; PRIOR FILING DATE: 1998-04-29  
111 ; PRIOR APPLICATION NUMBER: 60/083495  
112 ; PRIOR FILING DATE: 1998-04-29  
113 ; PRIOR APPLICATION NUMBER: 60/083496  
114 ; PRIOR FILING DATE: 1998-04-29  
115 ; PRIOR APPLICATION NUMBER: 60/083499  
116 ; PRIOR FILING DATE: 1998-04-29  
117 ; PRIOR APPLICATION NUMBER: 60/083545  
118 ; PRIOR FILING DATE: 1998-04-29  
119 ; PRIOR APPLICATION NUMBER: 60/083554  
120 ; PRIOR FILING DATE: 1998-04-29  
121 ; PRIOR APPLICATION NUMBER: 60/083558  
122 ; PRIOR FILING DATE: 1998-04-29  
123 ; PRIOR APPLICATION NUMBER: 60/083559  
124 ; PRIOR FILING DATE: 1998-04-29  
125 ; PRIOR APPLICATION NUMBER: 60/083500  
126 ; PRIOR FILING DATE: 1998-04-29  
127 ; PRIOR APPLICATION NUMBER: 60/083742  
128 ; PRIOR FILING DATE: 1998-04-30  
129 ; PRIOR APPLICATION NUMBER: 60/084366  
130 ; PRIOR FILING DATE: 1998-05-05  
131 ; PRIOR APPLICATION NUMBER: 60/084414  
132 ; PRIOR FILING DATE: 1998-05-06  
133 ; PRIOR APPLICATION NUMBER: 60/084441  
134 ; PRIOR FILING DATE: 1998-05-06  
135 ; PRIOR APPLICATION NUMBER: 60/084637  
136 ; PRIOR FILING DATE: 1998-05-07  
137 ; PRIOR APPLICATION NUMBER: 60/084639  
138 ; PRIOR FILING DATE: 1998-05-07  
139 ; PRIOR APPLICATION NUMBER: 60/084640  
140 ; PRIOR FILING DATE: 1998-05-07  
141 ; PRIOR APPLICATION NUMBER: 60/084598  
142 ; PRIOR FILING DATE: 1998-05-07  
143 ; PRIOR APPLICATION NUMBER: 60/084600  
144 ; PRIOR FILING DATE: 1998-05-07  
145 ; PRIOR APPLICATION NUMBER: 60/084627  
146 ; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/084643  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/085339  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085338  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085323  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085582  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085700  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085689  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085579  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085580  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085573  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085704  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7% Score 8; DB 9; Length 67;

Best Local Similarity 100.0%; Pred. No. 36; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159  
 Db 12 GLLLLALL 19

# RESULT 68

US-09-399-832A-85  
 ; Sequence 85, Application US/09999832A  
 ; Publication No. US20020192706A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Baker Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Ferrara, Napoleon  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Hillan, Kenneth J.  
 ; APPLICANT: Kijavini, Ivar J.  
 ; APPLICANT: Kuo, Sophia S.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James;  
 ; APPLICANT: Paoli, Nicholas P.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Shelton, David L.  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; FILE REFERENCE: P2630P1C53  
 ; CURRENT APPLICATION NUMBER: US/09/999,832A  
 ; CURRENT FILING DATE: 2001-10-24  
 ; PRIOR APPLICATION NUMBER: 09/918585  
 ; PRIOR FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/064249  
 ; PRIOR FILING DATE: 1997-11-03  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066364  
 ; PRIOR FILING DATE: 1997-11-21  
 ; PRIOR APPLICATION NUMBER: 60/077450  
 ; PRIOR FILING DATE: 1998-03-10  
 ; PRIOR APPLICATION NUMBER: 60/077632  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077641  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077649  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077791  
 ; PRIOR FILING DATE: 1998-03-12  
 ; PRIOR APPLICATION NUMBER: 60/078004  
 ; PRIOR FILING DATE: 1998-03-13  
 ; PRIOR APPLICATION NUMBER: 60/078886  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078936  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078939  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/079294  
 ; PRIOR FILING DATE: 1998-03-25  
 ; PRIOR APPLICATION NUMBER: 60/079656  
 ; PRIOR FILING DATE: 1998-03-26  
 ; PRIOR APPLICATION NUMBER: 60/079664  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079689  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079663  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079728  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079786  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079920  
 ; PRIOR FILING DATE: 1998-03-30  
 ; PRIOR APPLICATION NUMBER: 60/079923  
 ; PRIOR FILING DATE: 1998-03-30  
 ; PRIOR APPLICATION NUMBER: 60/080105  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080107  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080165  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080194  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080327  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080328  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080333  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080334  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/081070  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081049  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081071  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081195  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081203  
 ; PRIOR FILING DATE: 1998-04-09  
 ; PRIOR APPLICATION NUMBER: 60/081229  
 ; PRIOR FILING DATE: 1998-04-09



; PRIOR APPLICATION NUMBER: 60/081955  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081817  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081819  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081952  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081838  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/082568  
 ; PRIOR FILING DATE: 1998-04-21  
 ; PRIOR APPLICATION NUMBER: 60/082569  
 ; PRIOR FILING DATE: 1998-04-21  
 ; PRIOR APPLICATION NUMBER: 60/082704  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082804  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082700  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082797  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082796  
 ; PRIOR FILING DATE: 1998-04-23  
 ; PRIOR APPLICATION NUMBER: 60/083336  
 ; PRIOR FILING DATE: 1998-04-27  
 ; PRIOR APPLICATION NUMBER: 60/083322  
 ; PRIOR FILING DATE: 1998-04-28  
 ; PRIOR APPLICATION NUMBER: 60/083392  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083495  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083496  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083499  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083545  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083554  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083558  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083559  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083500  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083742  
 ; PRIOR FILING DATE: 1998-04-30  
 ; PRIOR APPLICATION NUMBER: 60/084366  
 ; PRIOR FILING DATE: 1998-05-05  
 ; PRIOR APPLICATION NUMBER: 60/084414  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: 60/084441  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: 60/084637  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084639  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084640  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084598  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084600  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084627  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084643  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/085339  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085338  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085323

; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085582  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085700  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085689  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085579  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085580  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085573  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085704  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 9; Length 67;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLMLALL 1159  
 Db 12 GLMLALL 19

RESULT 69

US-09-978-189-85  
 ; Sequence 85, Application US/09978189  
 ; Publication No. US20030004102A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Baker Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Baton, Dan  
 ; APPLICANT: Ferrara, Napoleon  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Hillan, Kenneth J.  
 ; APPLICANT: Kljavin, Ivar J.  
 ; APPLICANT: Kuo, Sophia S.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas F.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Shelton, David L.  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; FILE REFERENCE: P2630P1C7  
 ; CURRENT APPLICATION NUMBER: US/09/978,189  
 ; CURRENT FILING DATE: 2001-10-15  
 ; PRIOR APPLICATION NUMBER: 09/918585  
 ; PRIOR FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: 60/062250  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/064249  
 ; PRIOR FILING DATE: 1997-11-03  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066364  
 ; PRIOR FILING DATE: 1997-11-21

1 PRIOR APPLICATION NUMBER: 60/077450  
2 PRIOR FILING DATE: 1998-03-10  
3 PRIOR APPLICATION NUMBER: 60/077632  
4 PRIOR FILING DATE: 1998-03-11  
5 PRIOR APPLICATION NUMBER: 60/077641  
6 PRIOR FILING DATE: 1998-03-11  
7 PRIOR APPLICATION NUMBER: 60/077649  
8 PRIOR FILING DATE: 1998-03-11  
9 PRIOR APPLICATION NUMBER: 60/077791  
10 PRIOR FILING DATE: 1998-03-12  
11 PRIOR APPLICATION NUMBER: 60/078004  
12 PRIOR FILING DATE: 1998-03-13  
13 PRIOR APPLICATION NUMBER: 60/078886  
14 PRIOR FILING DATE: 1998-03-20  
15 PRIOR APPLICATION NUMBER: 60/078936  
16 PRIOR FILING DATE: 1998-03-20  
17 PRIOR APPLICATION NUMBER: 60/078910  
18 PRIOR FILING DATE: 1998-03-20  
19 PRIOR APPLICATION NUMBER: 60/078939  
20 PRIOR FILING DATE: 1998-03-20  
21 PRIOR APPLICATION NUMBER: 60/079294  
22 PRIOR FILING DATE: 1998-03-25  
23 PRIOR APPLICATION NUMBER: 60/079656  
24 PRIOR FILING DATE: 1998-03-26  
25 PRIOR APPLICATION NUMBER: 60/079664  
26 PRIOR FILING DATE: 1998-03-27  
27 PRIOR APPLICATION NUMBER: 60/079689  
28 PRIOR FILING DATE: 1998-03-27  
29 PRIOR APPLICATION NUMBER: 60/079663  
30 PRIOR FILING DATE: 1998-03-27  
31 PRIOR APPLICATION NUMBER: 60/079728  
32 PRIOR FILING DATE: 1998-03-27  
33 PRIOR APPLICATION NUMBER: 60/079786  
34 PRIOR FILING DATE: 1998-03-27  
35 PRIOR APPLICATION NUMBER: 60/079920  
36 PRIOR FILING DATE: 1998-03-30  
37 PRIOR APPLICATION NUMBER: 60/079923  
38 PRIOR FILING DATE: 1998-03-30  
39 PRIOR APPLICATION NUMBER: 60/080105  
40 PRIOR FILING DATE: 1998-03-31  
41 PRIOR APPLICATION NUMBER: 60/080107  
42 PRIOR FILING DATE: 1998-03-31  
43 PRIOR APPLICATION NUMBER: 60/080155  
44 PRIOR FILING DATE: 1998-03-31  
45 PRIOR APPLICATION NUMBER: 60/080194  
46 PRIOR FILING DATE: 1998-03-31  
47 PRIOR APPLICATION NUMBER: 60/080327  
48 PRIOR FILING DATE: 1998-04-01  
49 PRIOR APPLICATION NUMBER: 60/080328  
50 PRIOR FILING DATE: 1998-04-01  
51 PRIOR APPLICATION NUMBER: 60/080333  
52 PRIOR FILING DATE: 1998-04-01  
53 PRIOR APPLICATION NUMBER: 60/080334  
54 PRIOR FILING DATE: 1998-04-01  
55 PRIOR APPLICATION NUMBER: 60/081070  
56 PRIOR FILING DATE: 1998-04-08  
57 PRIOR APPLICATION NUMBER: 60/081049  
58 PRIOR FILING DATE: 1998-04-08  
59 PRIOR APPLICATION NUMBER: 60/081071  
60 PRIOR FILING DATE: 1998-04-08  
61 PRIOR APPLICATION NUMBER: 60/081195  
62 PRIOR FILING DATE: 1998-04-08  
63 PRIOR APPLICATION NUMBER: 60/081203  
64 PRIOR FILING DATE: 1998-04-09  
65 PRIOR APPLICATION NUMBER: 60/081229  
66 PRIOR FILING DATE: 1998-04-09  
67 PRIOR APPLICATION NUMBER: 60/081955  
68 PRIOR FILING DATE: 1998-04-15  
69 PRIOR APPLICATION NUMBER: 60/081817  
70 PRIOR FILING DATE: 1998-04-15  
71 PRIOR APPLICATION NUMBER: 60/081819  
72 PRIOR FILING DATE: 1998-04-15  
73 PRIOR APPLICATION NUMBER: 60/081952  
74 PRIOR FILING DATE: 1998-04-15  
75 PRIOR APPLICATION NUMBER: 60/081838  
76 PRIOR FILING DATE: 1998-04-15  
77 PRIOR APPLICATION NUMBER: 60/082568  
78 PRIOR FILING DATE: 1998-04-21  
79 PRIOR APPLICATION NUMBER: 60/082569  
80 PRIOR FILING DATE: 1998-04-21  
81 PRIOR APPLICATION NUMBER: 60/082704  
82 PRIOR FILING DATE: 1998-04-22  
83 PRIOR APPLICATION NUMBER: 60/082804  
84 PRIOR FILING DATE: 1998-04-22  
85 PRIOR APPLICATION NUMBER: 60/082700  
86 PRIOR FILING DATE: 1998-04-22  
87 PRIOR APPLICATION NUMBER: 60/082797  
88 PRIOR FILING DATE: 1998-04-22  
89 PRIOR APPLICATION NUMBER: 60/082796  
90 PRIOR FILING DATE: 1998-04-23  
91 PRIOR APPLICATION NUMBER: 60/083336  
92 PRIOR FILING DATE: 1998-04-27  
93 PRIOR APPLICATION NUMBER: 60/083322  
94 PRIOR FILING DATE: 1998-04-28  
95 PRIOR APPLICATION NUMBER: 60/083392  
96 PRIOR FILING DATE: 1998-04-29  
97 PRIOR APPLICATION NUMBER: 60/083495  
98 PRIOR FILING DATE: 1998-04-29  
99 PRIOR APPLICATION NUMBER: 60/083496  
100 PRIOR FILING DATE: 1998-04-29  
101 PRIOR APPLICATION NUMBER: 60/083499  
102 PRIOR FILING DATE: 1998-04-29  
103 PRIOR APPLICATION NUMBER: 60/083545  
104 PRIOR FILING DATE: 1998-04-29  
105 PRIOR APPLICATION NUMBER: 60/083554  
106 PRIOR FILING DATE: 1998-04-29  
107 PRIOR APPLICATION NUMBER: 60/083558  
108 PRIOR FILING DATE: 1998-04-29  
109 PRIOR APPLICATION NUMBER: 60/083559  
110 PRIOR FILING DATE: 1998-04-29  
111 PRIOR APPLICATION NUMBER: 60/083500  
112 PRIOR FILING DATE: 1998-04-29  
113 PRIOR APPLICATION NUMBER: 60/083742  
114 PRIOR FILING DATE: 1998-04-30  
115 PRIOR APPLICATION NUMBER: 60/084366  
116 PRIOR FILING DATE: 1998-05-05  
117 PRIOR APPLICATION NUMBER: 60/084414  
118 PRIOR FILING DATE: 1998-05-06  
119 PRIOR APPLICATION NUMBER: 60/084441  
120 PRIOR FILING DATE: 1998-05-06  
121 PRIOR APPLICATION NUMBER: 60/084637  
122 PRIOR FILING DATE: 1998-05-07  
123 PRIOR APPLICATION NUMBER: 60/084639  
124 PRIOR FILING DATE: 1998-05-07  
125 PRIOR APPLICATION NUMBER: 60/084640  
126 PRIOR FILING DATE: 1998-05-07  
127 PRIOR APPLICATION NUMBER: 60/084598  
128 PRIOR FILING DATE: 1998-05-07  
129 PRIOR APPLICATION NUMBER: 60/084600  
130 PRIOR FILING DATE: 1998-5-07  
131 PRIOR APPLICATION NUMBER: 60/084627  
132 PRIOR FILING DATE: 1998-05-07  
133 PRIOR APPLICATION NUMBER: 60/084643  
134 PRIOR FILING DATE: 1998-05-07  
135 PRIOR APPLICATION NUMBER: 60/085339  
136 PRIOR FILING DATE: 1998-05-13  
137 PRIOR APPLICATION NUMBER: 60/085338  
138 PRIOR FILING DATE: 1998-05-13  
139 PRIOR APPLICATION NUMBER: 60/085323  
140 PRIOR FILING DATE: 1998-05-13  
141 PRIOR APPLICATION NUMBER: 60/085582  
142 PRIOR FILING DATE: 1998-05-15  
143 PRIOR APPLICATION NUMBER: 60/085700  
144 PRIOR FILING DATE: 1998-05-15  
145 PRIOR APPLICATION NUMBER: 60/085689  
146 PRIOR FILING DATE: 1998-05-15

;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 10; Length 67;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159  
Db 12 GLLLLALL 19

## RESULT 70

US-09-978-608A-85  
; Sequence 85, Application US/09978608A  
; Publication No. US20030045462A1  
; GENERAL INFORMATION:

;; APPLICANT: Ashkenazi, Avi  
;; APPLICANT: Baker Kevin P.  
;; APPLICANT: Botstein, David  
;; APPLICANT: Desnovers, Luc  
;; APPLICANT: Eaton, Dan  
;; APPLICANT: Ferrara, Napoleon  
;; APPLICANT: Filvaroff, Ellen  
;; APPLICANT: Fong, Sherman  
;; APPLICANT: Gerber, Hanspeter  
;; APPLICANT: Gerritsen, Mary E.  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Grimaldi, J. Christopher  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Hillan, Kenneth J.  
;; APPLICANT: Kljavin, Ivar J.  
;; APPLICANT: Kuo, Sophia S.  
;; APPLICANT: Napier, Mary A.  
;; APPLICANT: Pan, James;  
;; APPLICANT: Paoni, Nicholas F.  
;; APPLICANT: Roy, Margaret Ann  
;; APPLICANT: Shelton, David L.  
;; APPLICANT: Stewart, Timothy A.  
;; APPLICANT: Tumas, Daniel  
;; APPLICANT: Williams, P. Mickey  
;; APPLICANT: Wood, William I.  
;; APPLICANT: Yoon, William I.  
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
;; FILE REFERENCE: P2630P1C22  
;; CURRENT APPLICATION NUMBER: US/09/978,608A  
;; CURRENT FILING DATE: 2001-10-16  
;; NUMBER OF SEQ ID NOS: 624  
;; Prior Application removed - See File Wrapper or Palm  
;; SEQ ID NO 85  
;; LENGTH: 67  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens

US-09-978-608A-85

Query Match 0.7%; Score 8; DB 10; Length 67;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159  
Db 12 GLLLLALL 19

## RESULT 71

US-09-978-585A-85  
; Sequence 85, Application US/09978585A  
; Publication No. US20030049633A1  
; GENERAL INFORMATION:

;; APPLICANT: Ashkenazi, Avi  
;; APPLICANT: Baker Kevin P.  
;; APPLICANT: Botstein, David  
;; APPLICANT: Desnovers, Luc  
;; APPLICANT: Eaton, Dan  
;; APPLICANT: Ferrara, Napoleon  
;; APPLICANT: Filvaroff, Ellen  
;; APPLICANT: Fong, Sherman  
;; APPLICANT: Gao, Wei-Qiang  
;; APPLICANT: Gerber, Hanspeter  
;; APPLICANT: Gerritsen, Mary E.  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Grimaldi, J. Christopher  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Hillan, Kenneth J.  
;; APPLICANT: Kljavin, Ivar J.  
;; APPLICANT: Kuo, Sophia S.  
;; APPLICANT: Napier, Mary A.  
;; APPLICANT: Pan, James;  
;; APPLICANT: Paoni, Nicholas F.  
;; APPLICANT: Roy, Margaret Ann  
;; APPLICANT: Shelton, David L.  
;; APPLICANT: Stewart, Timothy A.  
;; APPLICANT: Tumas, Daniel  
;; APPLICANT: Williams, P. Mickey  
;; APPLICANT: Wood, William I.  
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
;; FILE REFERENCE: P2630P1C15  
;; CURRENT APPLICATION NUMBER: US/09/978,585A  
;; CURRENT FILING DATE: 2001-10-16  
;; NUMBER OF SEQ ID NOS: 624  
;; Prior Application removed - See File Wrapper or Palm  
;; SEQ ID NO 85  
;; LENGTH: 67  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens

US-09-978-585A-85

Query Match 0.7%; Score 8; DB 10; Length 67;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159  
Db 12 GLLLLALL 19

## RESULT 72

US-09-978-191A-85  
; Sequence 85, Application US/09978191A  
; Publication No. US20030050239A1  
; GENERAL INFORMATION:

;; APPLICANT: Ashkenazi, Avi  
;; APPLICANT: Baker Kevin P.  
;; APPLICANT: Botstein, David  
;; APPLICANT: Desnovers, Luc  
;; APPLICANT: Eaton, Dan  
;; APPLICANT: Ferrara, Napoleon  
;; APPLICANT: Filvaroff, Ellen  
;; APPLICANT: Fong, Sherman  
;; APPLICANT: Gao, Wei-Qiang  
;; APPLICANT: Gerber, Hanspeter  
;; APPLICANT: Gerritsen, Mary E.  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Grimaldi, J. Christopher

APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James.  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630P1C4  
CURRENT APPLICATION NUMBER: US/09/978,191A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06

```

; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match      0.7%; Score 8; DB 10; Length 67;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 12 GLLLLALL 19

RESULT 73
US-09-978-403A-85
; Sequence 85, Application US/09978403A
; Publication No. US20030050240A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas P.

```

```

; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C17
; CURRENT APPLICATION NUMBER: US/09/978,403A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/085311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01

```

; PRIOR APPLICATION NUMBER: 60/080334  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/081070  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081049  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081071  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081195  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081203  
 ; PRIOR FILING DATE: 1998-04-09  
 ; PRIOR APPLICATION NUMBER: 60/081229  
 ; PRIOR FILING DATE: 1998-04-09  
 ; PRIOR APPLICATION NUMBER: 60/081955  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081817  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081819  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081952  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081838  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/082568  
 ; PRIOR FILING DATE: 1998-04-21  
 ; PRIOR APPLICATION NUMBER: 60/082569  
 ; PRIOR FILING DATE: 1998-04-21  
 ; PRIOR APPLICATION NUMBER: 60/082704  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082804  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082700  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082797  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082796  
 ; PRIOR FILING DATE: 1998-04-23  
 ; PRIOR APPLICATION NUMBER: 60/083336  
 ; PRIOR FILING DATE: 1998-04-27  
 ; PRIOR APPLICATION NUMBER: 60/083322  
 ; PRIOR FILING DATE: 1998-04-28  
 ; PRIOR APPLICATION NUMBER: 60/083392  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083495  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083496  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083499  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083545  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083554  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083558  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083559  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083500  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083742  
 ; PRIOR FILING DATE: 1998-04-30  
 ; PRIOR APPLICATION NUMBER: 60/084366  
 ; PRIOR FILING DATE: 1998-05-05  
 ; PRIOR APPLICATION NUMBER: 60/084414  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: 60/084441  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: 60/084637  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084639  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084640

; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084598  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084600  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084627  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084643  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/085339  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085338  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085323  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085582  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085700  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085689  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085579  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085580  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085573  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085704  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 10; Length 67;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLJLALL 1159  
 |||||  
 DB 12 GLLJLALL 19

RESULT 74  
 US-09-978-564A-85  
 ; Sequence 85, Application US/09978564A  
 ; Publication No. US20030050241A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Ferrara, Napoleon  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Geritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Hillan, Kenneth J.  
 ; APPLICANT: Kljavin, Ivar J.  
 ; APPLICANT: Kuo, Sophia S.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas F.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Shelton, David L.  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

1 TITLE OF INVENTION: Acids Encoding the Same  
2 FILE REFERENCE: P2630PIC25  
3 CURRENT APPLICATION NUMBER: US/09/978,564A  
4 CURRENT FILING DATE: 2001-10-16  
5 PRIOR APPLICATION NUMBER: 09/918585  
6 PRIOR FILING DATE: 2001-07-30  
7 PRIOR APPLICATION NUMBER: 60/062250  
8 PRIOR FILING DATE: 1997-10-17  
9 PRIOR APPLICATION NUMBER: 60/064249  
10 PRIOR FILING DATE: 1997-11-03  
11 PRIOR APPLICATION NUMBER: 60/065311  
12 PRIOR FILING DATE: 1997-11-13  
13 PRIOR APPLICATION NUMBER: 60/066364  
14 PRIOR FILING DATE: 1997-11-21  
15 PRIOR APPLICATION NUMBER: 60/074450  
16 PRIOR FILING DATE: 1998-03-10  
17 PRIOR APPLICATION NUMBER: 60/076732  
18 PRIOR FILING DATE: 1998-03-11  
19 PRIOR APPLICATION NUMBER: 60/077641  
20 PRIOR FILING DATE: 1998-03-11  
21 PRIOR APPLICATION NUMBER: 60/077649  
22 PRIOR FILING DATE: 1998-03-11  
23 PRIOR APPLICATION NUMBER: 60/077791  
24 PRIOR FILING DATE: 1998-03-12  
25 PRIOR APPLICATION NUMBER: 60/078004  
26 PRIOR FILING DATE: 1998-03-13  
27 PRIOR APPLICATION NUMBER: 60/078886  
28 PRIOR FILING DATE: 1998-03-20  
29 PRIOR APPLICATION NUMBER: 60/078936  
30 PRIOR FILING DATE: 1998-03-20  
31 PRIOR APPLICATION NUMBER: 60/078910  
32 PRIOR FILING DATE: 1998-03-20  
33 PRIOR APPLICATION NUMBER: 60/078939  
34 PRIOR FILING DATE: 1998-03-20  
35 PRIOR APPLICATION NUMBER: 60/079294  
36 PRIOR FILING DATE: 1998-03-25  
37 PRIOR APPLICATION NUMBER: 60/079656  
38 PRIOR FILING DATE: 1998-03-26  
39 PRIOR APPLICATION NUMBER: 60/079664  
40 PRIOR FILING DATE: 1998-03-27  
41 PRIOR APPLICATION NUMBER: 60/079689  
42 PRIOR FILING DATE: 1998-03-27  
43 PRIOR APPLICATION NUMBER: 60/079663  
44 PRIOR FILING DATE: 1998-03-27  
45 PRIOR APPLICATION NUMBER: 60/079728  
46 PRIOR FILING DATE: 1998-03-27  
47 PRIOR APPLICATION NUMBER: 60/079786  
48 PRIOR FILING DATE: 1998-03-27  
49 PRIOR APPLICATION NUMBER: 60/079920  
50 PRIOR FILING DATE: 1998-03-30  
51 PRIOR APPLICATION NUMBER: 60/079923  
52 PRIOR FILING DATE: 1998-03-30  
53 PRIOR APPLICATION NUMBER: 60/080105  
54 PRIOR FILING DATE: 1998-03-31  
55 PRIOR APPLICATION NUMBER: 60/080107  
56 PRIOR FILING DATE: 1998-03-31  
57 PRIOR APPLICATION NUMBER: 60/080165  
58 PRIOR FILING DATE: 1998-03-31  
59 PRIOR APPLICATION NUMBER: 60/080194  
60 PRIOR FILING DATE: 1998-03-31  
61 PRIOR APPLICATION NUMBER: 60/080327  
62 PRIOR FILING DATE: 1998-04-01  
63 PRIOR APPLICATION NUMBER: 60/080328  
64 PRIOR FILING DATE: 1998-04-01  
65 PRIOR APPLICATION NUMBER: 60/080333  
66 PRIOR FILING DATE: 1998-04-01  
67 PRIOR APPLICATION NUMBER: 60/080334  
68 PRIOR FILING DATE: 1998-04-01  
69 PRIOR APPLICATION NUMBER: 60/081070  
70 PRIOR FILING DATE: 1998-04-08  
71 PRIOR APPLICATION NUMBER: 60/081049  
72 PRIOR FILING DATE: 1998-04-08  
73 PRIOR APPLICATION NUMBER: 60/081071  
74 PRIOR FILING DATE: 1998-04-08  
75 PRIOR APPLICATION NUMBER: 60/081195  
76 PRIOR FILING DATE: 1998-04-08  
77 PRIOR APPLICATION NUMBER: 60/081203  
78 PRIOR FILING DATE: 1998-04-09  
79 PRIOR APPLICATION NUMBER: 60/081229  
80 PRIOR FILING DATE: 1998-04-09  
81 PRIOR APPLICATION NUMBER: 60/081955  
82 PRIOR FILING DATE: 1998-04-15  
83 PRIOR APPLICATION NUMBER: 60/081817  
84 PRIOR FILING DATE: 1998-04-15  
85 PRIOR APPLICATION NUMBER: 60/081819  
86 PRIOR FILING DATE: 1998-04-15  
87 PRIOR APPLICATION NUMBER: 60/081952  
88 PRIOR FILING DATE: 1998-04-15  
89 PRIOR APPLICATION NUMBER: 60/081838  
90 PRIOR FILING DATE: 1998-04-15  
91 PRIOR APPLICATION NUMBER: 60/082568  
92 PRIOR FILING DATE: 1998-04-21  
93 PRIOR APPLICATION NUMBER: 60/082569  
94 PRIOR FILING DATE: 1998-04-21  
95 PRIOR APPLICATION NUMBER: 60/082704  
96 PRIOR FILING DATE: 1998-04-22  
97 PRIOR APPLICATION NUMBER: 60/082804  
98 PRIOR FILING DATE: 1998-04-22  
99 PRIOR APPLICATION NUMBER: 60/082700  
100 PRIOR FILING DATE: 1998-04-22  
101 PRIOR APPLICATION NUMBER: 60/082797  
102 PRIOR FILING DATE: 1998-04-22  
103 PRIOR APPLICATION NUMBER: 60/082796  
104 PRIOR FILING DATE: 1998-04-23  
105 PRIOR APPLICATION NUMBER: 60/083336  
106 PRIOR FILING DATE: 1998-04-27  
107 PRIOR APPLICATION NUMBER: 60/083322  
108 PRIOR FILING DATE: 1998-04-28  
109 PRIOR APPLICATION NUMBER: 60/083392  
110 PRIOR FILING DATE: 1998-04-29  
111 PRIOR APPLICATION NUMBER: 60/083495  
112 PRIOR FILING DATE: 1998-04-29  
113 PRIOR APPLICATION NUMBER: 60/083496  
114 PRIOR FILING DATE: 1998-04-29  
115 PRIOR APPLICATION NUMBER: 60/083499  
116 PRIOR FILING DATE: 1998-04-29  
117 PRIOR APPLICATION NUMBER: 60/083545  
118 PRIOR FILING DATE: 1998-04-29  
119 PRIOR APPLICATION NUMBER: 60/083554  
120 PRIOR FILING DATE: 1998-04-29  
121 PRIOR APPLICATION NUMBER: 60/083558  
122 PRIOR FILING DATE: 1998-04-29  
123 PRIOR APPLICATION NUMBER: 60/083559  
124 PRIOR FILING DATE: 1998-04-29  
125 PRIOR APPLICATION NUMBER: 60/083500  
126 PRIOR FILING DATE: 1998-04-29  
127 PRIOR APPLICATION NUMBER: 60/083742  
128 PRIOR FILING DATE: 1998-04-30  
129 PRIOR APPLICATION NUMBER: 60/084366  
130 PRIOR FILING DATE: 1998-05-05  
131 PRIOR APPLICATION NUMBER: 60/084414  
132 PRIOR FILING DATE: 1998-05-06  
133 PRIOR APPLICATION NUMBER: 60/084441  
134 PRIOR FILING DATE: 1998-05-06  
135 PRIOR APPLICATION NUMBER: 60/084637  
136 PRIOR FILING DATE: 1998-05-07  
137 PRIOR APPLICATION NUMBER: 60/084639  
138 PRIOR FILING DATE: 1998-05-07  
139 PRIOR APPLICATION NUMBER: 60/084640  
140 PRIOR FILING DATE: 1998-05-07  
141 PRIOR APPLICATION NUMBER: 60/084598  
142 PRIOR FILING DATE: 1998-05-07  
143 PRIOR APPLICATION NUMBER: 60/084600  
144 PRIOR FILING DATE: 1998-05-07  
145 PRIOR APPLICATION NUMBER: 60/084627  
146 PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/084643  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/085339  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085338  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085323  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085582  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085700  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085689  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085579  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085580  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085573  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085704  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7; Score 8; DB 10; Length 67;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLALL 1159

Db 12 GLLALL 19

RESULT 75

US-09-999-833A-85  
; Sequence 85, Application US/09999833A

; Publication No. US20030054405A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnovers, Luc  
; APPLICANT: Baton, Dan  
; APPLICANT: Ferrara Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: F2630PIC65  
; CURRENT APPLICATION NUMBER: US/09/999,833A  
; CURRENT FILING DATE: 2001-10-24  
; PRIOR APPLICATION NUMBER: 09/918595  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/064249  
; PRIOR FILING DATE: 1997-11-03  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066364  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/077450  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: 60/077632  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077641  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077649  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077791  
; PRIOR FILING DATE: 1998-03-12  
; PRIOR APPLICATION NUMBER: 60/078004  
; PRIOR FILING DATE: 1998-03-13  
; PRIOR APPLICATION NUMBER: 60/078886  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078936  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078939  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 60/079664  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079663  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079786  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079920  
; PRIOR FILING DATE: 1998-03-30  
; PRIOR APPLICATION NUMBER: 60/079923  
; PRIOR FILING DATE: 1998-03-30  
; PRIOR APPLICATION NUMBER: 60/080105  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080107  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080165  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080194  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080327  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080328  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080333  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080334  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/081070  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081049  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081071  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081195  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081203  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081229  
; PRIOR FILING DATE: 1998-04-09



```

; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323

```

```

; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

```

Query Match 0.7%; Score 8; DB 10; Length 67;

Best Local Similarity 100.0%; Pred. No. 36; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1152 GLLLLALL 1159
Db 12 GLLLLALL 19

```

Search completed: June 24, 2004, 18:11:59  
Job time : 58 secs